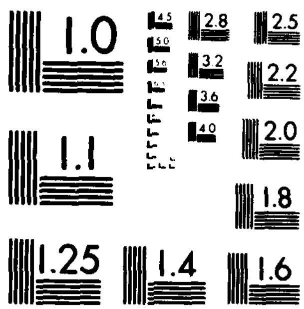
AD-A088 245 BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 12/1
APPROXIMATION SCHEMES FOR THE LINEAR-QUADRATIC OPTIMAL CONTROL --ETC(U)
MAR 80 K KUNISCH DAAG29-79-C-0161
UNCLASSIFIED AFOSR-TR-80-0577 NL

1 OF 1
AD
A088245

END
DATE
FILMED
9-80
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS 1963 A

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LEVEL II

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFOSR-TR-80-0577	2. GOVT ACCESSION NO.: AD-A088 245
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): APPROXIMATION SCHEMES FOR THE LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEM ASSOCIATED WITH DELAY-EQUATIONS.	5. TYPE OF REPORT & PERIOD COVERED: Interim report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Karl Kunisch	8. CONTRACT OR GRANT NUMBER(s): DAAG29-79-C-0161; AFOSR-76-3092
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Brown University, Lefschetz Center for Dynamical Systems, Providence, RI 02912	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332	12. REPORT DATE: March 1980
	13. NUMBER OF PAGES: 70
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public; release distribution unlimited.

DTIC ELECTE AUG 26 1980 S D E

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The linear regulator problem for delay equation is discussed. We propose a (theoretical) solution involving Riccati integral equations and them axiomatically discuss a general approximation scheme. The details are given for spline and averaging approximations.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 401834

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AFOSR-TR- 80-0577

APPROXIMATION SCHEMES FOR THE LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEM ASSOCIATED WITH DELAY-EQUATIONS+

by

Karl Kunisch
Institut fur Mathematik
Technische Universitat
Kopernikusgasse 24
A-8010 Graz, Austria

and

Division of Applied Mathematics
Lefschetz Center for Dynamical Systems
Brown University
Providence, R. I. 02912

March 1980

+This research was supported in part by the Air Force Office of Scientific Research under Contract [redacted]-AFOSR 76-3092[redacted] and in part by the U.S. Army Research Office, under ARO-DAAG 29-79-C-0161.

APPROXIMATION SCHEMES FOR THE LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEM ASSOCIATED WITH DELAY-EQUATIONS

Karl Kunisch

Abstract: The linear regulator problem for delay equation is discussed. We propose a (theoretical) solution involving Riccati integral equations and then axiomatically discuss a general approximation scheme. The details are given for spline-and averaging approximations.

1. Introduction and Notation

The problem of approximating delay-differential equations by sequences of either ordinary differential equations or algebraic equations has stimulated research for over fifteen years now. However, it was not until quite recently that convergence proofs in an operator-theoretic framework were given; see [1] and the references given there.

In this paper we address a specific problem of the above type, namely the approximation of the regulator problem of minimizing a quadratic cost-functional subject to a delay-or more generally a functional differential equation (FDE). This question also has attracted attention for quite some time . In [13,15] Ross and Flugge-Lotz and Solimon and Ray specify certain approximation schemes leaving open the question of convergence. In today's terminology their methods would be called averaging projections or linear interpolating spline scheme [1,8]. Not only does the question of approximation of the linear-quadratic control problem for (FDE) present difficulties, but the theoretic development of existence of solutions, deriving a feedback law and discussing an operator Riccati equation is challenging as well. Delfour treats these theoretic aspects in [4,5] and proves convergence of the averaging scheme, discretizing space and time variables. We also refer to [5] as a reference on the literature to the linear-quadratic optimal control problem for (FDE) up to (1977).

In the present paper we develop a general theory for the above mentioned problem, which we subsequently apply to the spline and averaging approximation schemes. The theoretical aspects are greatly facilitated by a recent paper of Gibson [6] in which an abstract linear-quadratic optimal control problem is treated in a general Hilbert space; it was observed that the Riccati-operators satisfy two (almost) equivalent Riccati integral equations, one of which coincides with the one used by Delfour in [5], the other one ((2.17) in this paper), although implicitly present, was not dealt with in [5]. It should be noted that in our presentation the treatment of the original problem (as opposed to the approximating ones) is based solely on integral equations. A second important feature is that we avoid using the infinitesimal generator of the adjoint of the solution semigroup associated with the (FDE). All the estimates depend heavily on the fact that even in the abstract formulation of the (FDE) (see (2.4)), the control term enters only as an operator with finite dimensional range.

Many of the technicalities here arise from the fact that we intend to not only prove convergence of optimal controls, trajectories, payoffs etc., but also want to give some error bounds. This leads to an essential difficulty which is described at length in Remark 2.1.

The paper is organized in the following way. Section 2 contains the statement of the problem and its (theoretical) solution. Then a sequence of approximating problems is specified and the convergence results are stated, leaving technical proofs to Section 5. In Section 3, we first show how the results of Section 2 can be used for spline approximation schemes. For linear and cubic splines we give all the details, demonstrating convergence of the linear spline scheme and quadratic convergence on certain subspaces of the cubic spline scheme. Averaging projection schemes are discussed in Section 4; the approximating equations in this case turn out to coincide with those proposed in [2], [13] and [15].

Most of the notation that is used throughout the paper is quite standard. For a closed interval $I \subset (-\infty,\infty)$, a Banach space X with norm $|\cdot|_X$ and $p \geq 1$, the equivalence class of measurable functions $x: I \to X$ with $\int_I |x(s)|_X^p ds < \infty$ is denoted by $L^p(I;X)$. $|\cdot|_{L^p(I;X)}$ or simply $|\cdot|_{L^p}$ is the notation for the usual norm in $L^p(I;X)$. The space of continuous functions on I with values in X endowed with the supremum norm is denoted by $C(I;X)$ and $C^k(I;X)$, $k = 1,2,\ldots$ stands for the space of X-valued continuous functions which possess k continuous derivatives on I. $W^{k,2}(I;X)$, $k = 1,2,\ldots$ is the space of (k-1)-times continuously differentiable functions whose (k-1)-st derivative is absolutely continuous with derivative in $L^2(I;X)$; $|\cdot|_{W^{k,2}(I;X)}$ denotes any one of the commonly employed $W^{k,2}$-norms. The space of all essentially bounded and strongly measurable functions from I to X is denoted by $\mathscr{B}_\infty(I;X)$. In the special case of $I = [-r,0]$, $0 < r < \infty$ and $X = \mathbb{R}^n$ we shall abbreviate the notation of the function spaces by $L^2, C^k, W^{k,2}$, etc.

For Banach spaces X and Y, the set of all bounded linear operators from X to Y is denoted by $\mathscr{L}(X,Y)$ and for $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^m)$ we simply write $\mathbb{R}^{n\times m}$. For $A \in \mathscr{L}(X,Y)$ the strong operator-norm is denoted by $||A||_{\mathscr{L}(X,Y)}$. A^* stands for the Hilbert space-adjoint of an operator A from a Hilbert space H to H. $\mathbb{R}^n$ is endowed with the euclidean norm $|\cdot|_{\mathbb{R}^n}$ and $(\cdot,\cdot)_{\mathbb{R}^n}$ stands for the usual inner product in $\mathbb{R}^n$. For elements in $\mathbb{R}^{n\times m}$ we use the spectral norm. Wherever the contents permits we drop the subscript of a norm, simply using $|\cdot|$ for the norm of elements of a Banach space and $||\cdot||$ for that of operators between Banach spaces.

The state space of our presentation will be $\mathbb{R}^n \times L^2(-r,0;\mathbb{R}^n)$ with the norm

$$|(\eta,\phi)| = \left(|\eta|_{\mathbb{R}^n} + \int_{-r}^{0} \rho(s)|\phi(s)|^2 ds\right)^{1/2}$$

where the weighting function $\rho\colon [-r,0] \to \mathbb{R}$ is a piecewise continuous and positive function. We denote by Z (or Z_ρ where necessary) the space $\mathbb{R}^n \times L^2(-r,0;\mathbb{R}^n)$ together with the weighted norm. The symbol $\langle\cdot,\cdot\rangle$ stands for the natural inner product in Z and P_1, P_2 denote the projections of Z onto its first and second components respectively. $\mathscr{C}^k$ and $\mathscr{W}^{k,2}$ stand for subspaces of Z given by $\{(\phi(0),\phi)|\ \phi \in C^k\}$ and $\{(\phi(0),\phi)|\ \phi \in W^{k,2}\}$ respectively.

A family $V(t,s)$ of operators in $\mathscr{L}(Z,Z)$ with $t_0 \le s \le t \le t^*$ is called evolution operator if $V(s,s)z = z$, if $V(t,s)z = V(t,r)V(r,s)z$ and if $t \to V(t,s)z$ is continuous for all $z \in Z$ and $t_0 \le s \le r \le t \le t^*$. The derivative of a function x is denoted by $\dot{x}$ or also x', and, finally, for $x\colon [-r,\alpha) \to \mathbb{R}^n$, $\alpha > 0$, the symbol x_t, $0 \le t \le \alpha$ stands for the function $[-r,0] \to X$ given by $x_t(s) = x(t+s)$ for $s \in [-r,0]$.

__Acknowledgement__:

The author would like to thank Professor H. T. Banks for encouragement to work on this problem and for various stimulating discussions.

2. Approximation of the Linear-Quadratic Control Problem

For $(\eta,\phi) = z \in Z$ and $(t_0,t^*) \in \mathbb{R} \times \mathbb{R}$ we consider the functional differential equation (FDE)

$$\text{(2.1)} \qquad \begin{cases} \dot{x}(t) = L(t,x_t) + f(t), \text{ for } t_0 \leq s \leq t \leq t^* \\ x(s) = \eta,\ x_s = \phi \end{cases}$$

where

$$\text{(2.2)} \qquad L(t,\phi) = \sum_{i=0}^{\ell} A_i(t)\phi(-r_i) + \int_{-r}^{0} A_{-1}(t,s)\phi(s)ds$$

Here we let $0 = r_0 < r_1 < \ldots < r_\ell = r$ and the matrix-valued functions A_i, for $i = -1,\ldots,\ell$ are considered as operators in $A_i \in C(t_0,t^*;\mathbb{R}^{n\times n})$, for $i = 0,\ldots,\ell$. and $A_{-1} \in C(t_0,t^*;L^2(-r,0;\mathbb{R}^{n\times n}))$ respectively, and $f \in L^2(t_0,t^*;\mathbb{R}^n)$.

We also need to restrict our attention to the homogeneous problem

$$\text{(2.3)} \qquad \begin{cases} \dot{x}(t) = L(t,x_t), \quad \text{for } t_0 \leq s \leq t \leq t^* \\ x(s) = \eta,\ x_s = \phi. \end{cases}$$

It is quite well known [8] that solutions to (2.1) and (2.3) exist and that they do not depend on the representation of an equivalence class $\phi \in L^2$. We shall denote the solutions by $x(\cdot,s;z,f)$ and $x(\cdot,s;z)$ respectively, dropping arguments if the context permits us to do so. Let $T(t,s)\colon Z \to Z$ be the solution operator associated with (2.3), i.e.

$$T(t,s)z = ((x(t,s;z),x_t(\cdot,s;z)) \quad \text{for} \quad t_0 \le s \le t \le t^*.$$

Then $T(t,s)$ is an evolution operator on $\Delta = \{(t,s) | \ t_0 \le s \le t \le t^*\}$.

In the next lemma the weighting function for the norm of Z is chosen identically 1.

Lemma 2.1. Exponential bounds on $T(t,s)$ are given by

$$|T(t,s)z|_{Z_1} \le Me^{\omega(t-s)}|z|_{Z_1} \quad \text{for} \quad (t,s) \in \Delta,$$

where

$$M = \left(1 + \sum_{i=1}^{\ell} \sup_{t\in[t_0,t^*]} ||A_i(t)||\right)^{1/2}$$

and

$$\omega = M^2 + \sup_{t\in[t_0,t^*]} ||A_{-1}(t,\cdot)||,$$

with $A_{-1}(t,\cdot)$ considered as an element in $L^2(-r,0;\mathbb{R}^{n\times n})$.

For the proof see [12, Theorems 2.1 and 3.5].

We return to (2.1) and recall the following variation of constants formula.

Lemma 2.2. If for $z_0 \in Z$ we define $z(t,s;z_0) \in Z$ by

$$z(t,s;z_0) = (x(t,s;z_0,f),\ x_t(\cdot,s;z_0,f)) \quad \text{for} \quad (t,s) \in \Delta$$

then

$$(2.4) \quad z(t,s;z_0) = T(t,s)z_0 + \int_s^t T(t,\sigma)(f(\sigma),0)d\sigma, \quad \text{for} \quad (t,s) \in \Delta.$$

This result is proved in [12] and in the autonomous case it also follows trivially from [1], [3], [9].

In this paper, we shall consider the following optimal control problem:

$$(\mathrm{P}) \begin{cases} \text{Find } u \in L^2(t_0,t^*;\mathbb{R}^m) \text{ which minimizes} \\ J(t_0,\eta,\phi,u) = (Fx(t^*),x(t^*))_{\mathbb{R}^n} + \int_{t_0}^{t^*} (D(t)x(t),x(t))_{\mathbb{R}^n} dt \\ \qquad + \int_{t_0}^{t^*} (C(t)u(t),u(t))_{\mathbb{R}^m} dt \\ \text{subject to} \\ \dot{x}(t) = L(t,x_t) + B(t)u(t), \quad t_0 \le t \le t^* \\ x(t_0) = \eta,\ x_{t_0} = \phi, \text{ where } (\eta,\phi) \in Z \text{ and } t_0,t^* \text{ are given.} \end{cases}$$

In the notation of the cost functional J we let $x(t)$ stand for $x(t,t_0;\eta,\phi,B(t)u(t))$. The assumptions on F,D,C and B are the following:

$$(2.5)\quad \begin{cases} F \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n), \text{ selfadjoint, nonnegative}, \\ D \in \mathscr{B}_\infty(t_0, t^*; \mathbb{R}^{n\times n}), \text{ selfadjoint, nonnegative}, \\ C \in \mathscr{B}_\infty(t_0, t^*, \mathbb{R}^{m\times m}), \text{ selfadjoint, } C(t) \geq c > 0 \\ \qquad \text{for some } c > 0 \text{ and almost all } t, \\ B \in \mathscr{B}_\infty(t_0, t^*; \mathbb{R}^{m\times n}). \end{cases}$$

For the presentation of the approximation results we choose a sequence of closed linear subspaces $\{Z^N\}_{N=1}$ of Z and orthogonal projections

$$P^N: Z \to Z^N, \qquad \text{for } N = 1,2,\ldots\ .$$

We shall also use the operator $Q_0: \mathbb{R}^n \to Z$ given by

$$Q_0\eta = (\eta, 0).$$

Of course, Q_0 can be represented as an $n \times n$ - Z-valued matrix by

$$Q_0 = \begin{pmatrix} (1,0) & & 0 \\ & \ddots & \\ 0 & & (1,\ ) \end{pmatrix}$$

where 0 stands for the zero-element in Z. In general, we shall not distinguish between the operator Q_0 and its representation. With this notation (2.4) can be written $z(t,s;z_0) = T(t,s)z_0 + \int_s^t T(t,\sigma)Q_0 f(\sigma)d\sigma$. Motivated by earlier work on approximation of (FDE) [1,3,8] we may impose the following hypotheses:

(H1) There exists a family of evolution operators $T^N(t,s): Z \to Z$, for $N = 1,2,\ldots$ and $(t,s) \in \Delta$ such that

(i) $||T^N(t,s)|| \leq \overline{M}e^{\overline{\omega}(t-s)}$ for some $\overline{M} > 0$, $\overline{\omega} \in \mathbb{R}$,

(ii) $T^N(t,s)Z^N \subset Z^N$ for all $(t,s) \in \Delta$,

(iii) there exists a real-valued function $\overline{\rho}$ such that

$$|T(t,s)z - T^N(t,s)z| \leq \overline{\rho}(N,z).$$

Of course in the examples that we have in mind $\overline{\rho}$ will tend to 0 at a certain rate as N goes to ∞; the dependence of $\overline{\rho}$ on z will also indicate possible dependence on derivatives of z (compare Section 3).

(H2) $\lim\limits_{N\to\infty} P^N z = z$ for all $z \in Z$.

To get estimates on the rate of convergence we need to introduce a family of operators $Q^N: \mathbb{R}^n \to Z$, which act as "smoothing operators" for Q_0.

(H3) There exists a sequence of operators $Q^N: \mathbb{R}^n \to Z$, $N = 1,2,\ldots,$ such that

(i) $Q^N\mathbb{R}^n \subset Z^N$

(ii) $||Q^N - Q_0||_{\mathscr{L}(\mathbb{R}^n;Z)} \leq \rho_Q(N)$ for some real-valued function ρ_Q

(iii) $||Q^N||_{\mathscr{L}(\mathbb{R}^n;Z)} \leq q$ for some $q \geq 1$, independent of N.

Throughout this section we assume (H1)-(H3) to hold. A possible candidate for Q^N is the matrix whose columns are the orthogonal projections of the columns of Q_0 onto Z^N. Notice that (H3)(i) implies that there exist a matrix $Q_0^N \in \mathscr{L}(\mathbb{R}^n,\mathbb{R}^n)$ and a function-valued matrix $Q_1^N \in L^2(-r,0;\mathbb{R}^{n\times n})$ such that $((Q_0^N)_j,(Q_1^N)_j) \in Z^N$ for $j = 1,\ldots,n$, where $(E)_j$ stands for the j^{th} column of a matrix E. In the examples that we have in mind Q^N can always be chosen as a diagonal matrix, with diagonal elements in $\mathbb{R} \times L^2(-r,0;\mathbb{R})$ approximating $(1,0) \in \mathbb{R} \times L^2(-r,0;\mathbb{R})$. The need for introducing the family Q^N to obtain estimates on the rate of convergence will become apparent from the analysis below. The underlying problem, however, can be explained for real-valued functions on $[-1,0]$.

Remark 2.1. To demonstrate the need for introducing the family of operators Q^N let g_0: $[-1,0] \in \mathbb{R}$ be given by

$$g_0(s) = \begin{cases} 1 & \text{for } s = 0 \\ 0 & \text{for } s \in [-1,0). \end{cases}$$

It is not hard to find a sequence of functions g_N: $[-1,0] \to \mathbb{R}$, such that (α) $g_N(0) = 1$, (β) $|g_N|_{L^2} = O(\frac{1}{N^\rho})$ for some $\rho > 0$, (γ) $g_N \in W^{1,2}(-1,0;\mathbb{R})$ and (δ) $|\dot{g}_N|_{L^1} \leq M_1$ for some M_1 independent of N. In fact, we may take

$$g_N(t) = \begin{cases} Nt + 1 & \text{for} \quad t \in [-\frac{1}{N}, 0] \\ 0 & \text{otherwise} \end{cases} ;$$

however, $|\dot{g}_n|_{L^2}$ diverges like $\sqrt{N}$. For functions in $W^{2,2}(-1,0;\mathbb{R})$ we analyze the question more precisely: there exists no family of functions $\{g_N\}$ such that

$$(2.7) \quad \begin{cases} \text{(i)} & \lim_N g_N(0) = 1, \text{ for all } N \\ \text{(ii)} & \lim_N |g_N|_{L^2} = 0, \text{ for some } \rho \geq 0 \\ \text{(iii)} & g_N \in W^{2,2}(-1,0;\mathbb{R}) \\ \text{(iv)} & |\dot{g}_N|_{L^2} \leq M_1 \text{ and } |\ddot{g}_N|_{L^2} \leq M_2, \text{ both uniformly in } N. \end{cases}$$

Proof. Assuming (i)-(iii), we argue that (iv) cannot hold. We first show that $\lim_N g_N(-1) = 0$. For suppose there exists a subsequence, again denoted by g_N such that $g_N(-1) \geq \alpha > 0$, for all N. (The case $\alpha < 0$ is treated similarly.) Then

$$g_N(\varepsilon-1) = g_N(-1) + \int_{-1}^{\varepsilon-1} \dot{g}_N(s)ds \geq \alpha - \sqrt{\varepsilon}\, M_1.$$

Therefore, there exists $\varepsilon_0 > 0$ and $\bar{\alpha} > 0$ such that $g_N(\varepsilon-1) \geq \bar{\alpha} > 0$ for all N and $\varepsilon \in [0,\varepsilon_0]$. This contradicts (ii). Next, we verify that $\lim_N \dot{g}_N(-1) = 0$. If not, there exists a subsequence, again denoted by g_N such that $\dot{g}_N(-1) \geq \tilde{\alpha} > 0$, for all N; (the case $\tilde{\alpha} < 0$ is treated similarly). Then

$$g_N(\varepsilon-1) = g_N(-1) + \varepsilon\dot{g}_N(-1) + \int_{-1}^{\varepsilon-1} (\varepsilon-1-s)\ddot{g}_N(s)ds \geq g_N(-1) + \varepsilon(\tilde{\alpha} - M_2\sqrt{\tfrac{\varepsilon}{3}}),$$

so that there exist constants $\tilde{\varepsilon}_0 > 0$ and $k > 0$ such that $g_N(\varepsilon-1) \geq g_N(-1) + \varepsilon k$, for all N and $\varepsilon \in [0,\tilde{\varepsilon}_0]$, which again contradicts (ii). In a similar way one can show that $\lim\limits_N \dot{g}_N(0) = \infty$. Since the left-hand-side in the next estimate tends to ∞

$$|\dot{g}_N(0) - \dot{g}_N(-1)| \leq \left(\int_{-1}^{0} |\ddot{g}_N(s)|^2 ds\right)^{1/2} ,$$

we see that (iv) is violated and hence the proof of the above claim is completed.

There is yet another way of considering properties (i)-(iv), interesting from the point of view of spline analysis. We let $\tilde{s}_N \in W^{2,2}(-1,0;\mathbb{R})$ denote the unique cubic Hermite spline function given by

$$(2.8)\qquad \begin{cases} \tilde{s}_N'(t_j^N) = \tilde{s}_N(t_j^N) = 0 \quad \text{for} \quad j = 2,\ldots,N \\ \tilde{s}_N(t_1^N) = 0, \quad \tilde{s}'(t_1^N) = \beta, \\ \tilde{s}_N(0) = 1, \quad \tilde{s}_N'(0) = \alpha, \end{cases}$$

for a partition $t_j^N = -\frac{j}{N}$, $j = 0,\ldots,N$ of $[-1,0]$. A simple calculation shows that

$$\tilde{s}_N(t) = \begin{cases} (-2N^3 + (\alpha+\beta)N^2)t^3 + (2\alpha N + \beta N - 3N^2)t^2 + \alpha t + 1 & \text{for } t \in [t_1^N,0] \\ 0 & \text{otherwise.} \end{cases}$$

We recall that the variational problem of finding the function $v \in W^{2,2}(-1,0;\mathbb{R})$ satisfying (2.8) and minimizing $|\ddot{v}|_{L^2}$ is exactly the cubic Hermite spline $\tilde{s}_N$; but $|\tilde{s}_N''|_{L^2}$ diverges like $N^{3/2}$. Of course, a similar negative result can be shown for cubic spline functions.

To relate the above observations to the operator Q_0 we suppose $r = n = 1$ and $Z^N \subset \{(\phi(0),\phi) \mid \phi \in W^{2,2}(-1,0;\mathbb{R})\}$. Then we have demonstrated that there does <u>not</u> exist a sequence of operators $Q^N: \mathbb{R} \to Z^N$ such that $\lim_N ||Q^N - Q_0||_{\mathscr{L}(\mathbb{R},Z)} = 0$ and such that $||d(P_2 Q^N)||_{\mathscr{L}(\mathbb{R},L^2(-1,0;\mathbb{R}))}$ and $||d^2(P_2 Q^N)||_{\mathscr{L}(\mathbb{R},L^2(-1,0;\mathbb{R}))}$ are uniformly bounded in N; here d denotes the differentiation operator.

To explain the significance of the above negative result, we recall that in order to get good convergence results in spline analysis, the (L^2-and Tschebyscheff-norm of the) derivatives of the approximated function play an essential role [11,14]. In the next section we shall apply the general result of this section to specific spline approximation schemes, and it is no surprise that again the convergence of $T^N(t,s)z$ to $T(t,s)z$ depends on the smoothness of z, (see [3]). A brief look at (2.4) indicates that this will cause severe difficulties, since under the integral the operator $T(t,s)$ always acts on a discontinuous function. The special form of the integral will help to get the convergence result. But for estimates of the rate of convergence of control, state, payoff and Riccati operator, certain uniformities in the convergence of Q^N to Q_0 would be needed, which for cubic spline approximations turn out to be those given in (2.7).

To get better results than just convergence, in spite of the above difficulties, we shall use the following simple technique, which we explain by using g_0: For some desired accurracy ε determine a function g_N (Q^N or $P^N Q_0$ later on in this paper) such that $g_N(0) = 1$ and $|g_N|_{L^2} \leq \varepsilon$. Moreover, g_N will be chosen in such a way that it suits our smoothness requirements for the specific situation and such that there exists a sequence of functions g_N^μ, whose convergence to g_N is of a desired rate. This ends Remark 2.1.

We return to the development of the theory begun prior to this long remark and aim at an "abstract" formulation in the space Z of problem (P). We shall need the operators $\mathscr{B}$ and $\mathscr{B}^N \in \mathscr{B}_\infty(t_0,t^*;\mathscr{L}(\mathbb{R}^m,Z))$, $\mathscr{D}$ and $\mathscr{D}^N \in \mathscr{B}_\infty(t_0,t^*;\mathscr{L}(Z,Z))$ and $\mathscr{F}$ and $\mathscr{F}^N \in \mathscr{L}(Z,Z)$ given by

$$\mathscr{B}(t) = Q_0B(t) \quad \text{and} \quad \mathscr{B}^N(t) = Q^NB(t),$$

$$\mathscr{F}(\eta,\phi) = (F(\eta),0) \quad \text{for} \quad (\eta,\phi) \in Z \quad \text{and} \quad F^N = P^N\mathscr{F}P^N,$$

$$\mathscr{D}(t)(\eta,\phi) = (D(t)(\eta),0) \quad \text{for} \quad (\eta,\phi) \in Z \quad \text{and} \quad \mathscr{D}^N(t) = P^N\mathscr{D}(t)P^N.$$

<u>Lemma 2.3</u>. The operators $\mathscr{B}$ and $\mathscr{B}^N, \mathscr{D}$ and $\mathscr{D}^N, \mathscr{F}$ and F^N satisfy the same properties as B, D and F in (2.5) respectively. Moreover,

$$\mathscr{B}^*(t)(\eta,\phi) = B^*(t)\eta, \text{ for } (\eta,\phi) \in Z \text{ and } t_0 \le t \le t^*,$$

$$(\mathscr{B}^N)^*(t)(\eta,\phi) = B^*(t)(Q_0^N)^*\eta + \int_{-r}^0 B^*(t)(Q_1^N)^*(s)\phi(s)ds,$$

and

$$(\mathscr{B}^N)^* \text{ and } \mathscr{B}^* \in \mathscr{B}_\infty(t_0,t^*;\mathscr{L}(Z,\mathbb{R}^m)).$$

<u>Proof</u>. We shall only verify the representation of $(\mathscr{B}^N)^*$. So let $v \in \mathbb{R}^m$, and $(\eta,\phi) \in Z$ be arbitrary, then

$$\langle \mathscr{B}^N v,(\eta,\phi)\rangle = (Q_0^N B(t)v,\eta)_{\mathbb{R}^n} + \int_{-r}^{0} (Q_1^N(s)B(t)v,\phi(s))_{\mathbb{R}^m} ds$$

$$= (v,B^*(t)(Q_0^N)^*\eta)_{\mathbb{R}^m} + \int_{-r}^{0} (v,B^*(t)(Q_1^N)^*(s)\phi(s))_{\mathbb{R}^m} ds$$

$$= (v,B^*(t)[(Q_0^N)^*\eta + \int_{-r}^{0} (Q_1^N)^*(s)\phi(s)])_{\mathbb{R}^m}.$$

Next we introduce the family of approximating optimal control problems in which the original problem $(\mathscr{P})$ is imbedded. Let

$$T^0(t,s) = T(t,s),\ P^0 = I,\ \mathscr{B}^0(t) = \mathscr{B}(t), Q^0 = Q_0,$$
$$\mathscr{D}^0 = \mathscr{D} \text{ and } F^0 = \mathscr{F},$$

and consider

$$(\mathscr{P}^{N,\mu}) \quad (2.9) \quad \begin{cases} \text{For } t_0 \in \mathbb{R},\ t^* \in \mathbb{R} \text{ and } z \in Z \text{ given, minimize} \\ J(t_0,P^{N+\mu}z,u) = \langle F^{N+\mu}z^{N,\mu}(t^*),z^{N,\mu}(t^*)\rangle + \\ \qquad + \int_{t_0}^{t^*} (\langle \mathscr{D}^{N+\mu}(t)z^{N,\mu}(t),z^{N,\mu}(t)\rangle dt + (C(t)u(t),u(t)))dt \\ \text{over } u \in L^2(t_0,t^*;\mathbb{R}^m) \text{ subject to} \\ z^{N,\mu}(t) = T^{N+\mu}(t,t_0)P^{N+\mu}z + \int_{t_0}^{t} T^{N+\mu}(t,\eta)\mathscr{B}^N(\eta)u(\eta)d\eta \\ \qquad\qquad \text{for } t_0 \le t \le t^*. \end{cases}$$

For $\mu = N = 0$ and by Lemma 2.2 we have $(\mathscr{P}^{0,0}) = (\mathscr{P})$. We first address ourselves to solving $(\mathcal{P}^{N,\mu})$, $N,\mu = 0,1,\ldots$ and to the question of the behavior of the solution of $(\mathcal{P}^{N,\mu})$ as $N,\mu \to \infty$. Problems $(\mathcal{P}^{N,\mu})$ are special cases of the linear quadratic optimal control problem considered in [6]. Applying this result we outline the solution of $(\mathcal{P}^{N,\mu})$.

Under the assumptions on F,D and C the unique optimal controls are the solutions of

$$J'(t_0,P^{N+\mu}z(t_0),u)v = 0, \quad \text{for all} \quad v \in L^2(t_0,t^*;\mathbb{R}^m),$$

where

$$J'(t_0,P^{N+\mu}z(t_0),u)v$$

denotes the Fréchet-derivative of J at u in the direction v. Therefore, after some calculations one finds that the optimal controls $\tilde{u}^{N,\mu}$ are given by

$$\text{(2.10)} \qquad \tilde{u}^{N,\mu}(t) = -((V_{t_0}^{N,\mu})^{-1}W_{t_0}^{N,\mu}z)(t) \quad \text{a.e. in} \quad [t_0,t^*]$$

where

$$V_{t_0}^{N,\mu} \in \mathcal{L}(L^2(t_0,t^*;\mathbb{R}^m),\ L^2(t_0,t^*;\mathbb{R}^m))$$

and

$$W_{t_0}^{N,\mu} \in \mathcal{L}(Z,L^2(t_0,t^*;\mathbb{R}^m))$$

and

$$(2.11)\qquad V_{t_0}^{N,\mu} = C + (\mathscr{B}^N)^*(\mathscr{T}_{t_0}^{N+\mu})^*\mathscr{Q}^{N+\mu}\mathscr{T}_{t_0}^{N+\mu}\mathscr{B}^N + (\mathscr{B}^N)^*(\mathscr{F}_{t_0}^{N+\mu})^*F^{N+\mu}\,\mathscr{F}_{t_0}^{N+\mu}\mathscr{B}^N.$$

$$(2.12)\qquad W_{t_0}^{N,\mu} = (\mathscr{B}^N)^*(\mathscr{T}_{t_0}^{N+\mu})^*\mathscr{Q}^{N+\mu}T_{t_0}^{N+\mu} + (\mathscr{B}^N)^*(\mathscr{F}_{t_0}^{N+\mu})^*F^{N+\mu}T^{N+\mu}T(t^*,t_0).$$

Here

$$\mathscr{T}_{t_0}^{N+\mu} \in \mathscr{L}(L^2(t_0,t^*;Z),\ L^2(t_0,t^*;Z))$$

$$\mathscr{F}_{t_0}^{N+\mu} \in \mathscr{L}(L^2(t_0,t^*;Z),Z)$$

$$T_{t_0}^{N+\mu} \in \mathscr{L}(Z,L^2(t_0,t^*;Z))$$

are defined by

$$(\mathscr{T}^{N+\mu}\phi)(s) = \int_{t_0}^{s} T^{N+\mu}(s,\eta)\phi(\eta)d\eta, \quad \text{for} \quad \phi \in L^2(t_0,t^*;Z)$$

$$(\mathscr{F}_{t_0}^{N+\mu}\phi) = (\mathscr{T}_{t_0}^{N+\mu}\phi)(t^*),$$

$$((\mathscr{T}^{N+\mu})^*\phi)(t) = \int_{t}^{t^*} (T^{N+\mu})^*(\eta,t)\phi(\eta)d\eta,$$

$$((\mathscr{F}_{t_0}^{N+\mu})^*\tilde{z})(t) = (T^{N+\mu})^*(t^*,t)\tilde{z}, \quad \text{for} \quad \tilde{z} \in Z$$

$$(T_{t_0}^{N+\mu}\tilde{z})(t) = T^{N+\mu}(t,t_0)\tilde{z}.$$

Consider for a moment the optimal control problems $(\mathscr{P}^{N,\mu})$ with $J(t_0,P^{N+\mu}z,u)$ replaced by $J(s,P^{N+\mu}z^{N,\mu}(s),u)$, $t_0 \le s \le t^*$; letting

$\tilde{u}_s^{N,\mu}$ denote the corresponding optimal control, then it is clear from the above that

$$(2.13)\qquad (\tilde{u}_s^{N,\mu})(t) = -((V_s^{N,\mu})^{-1}W_s^{N,\mu}P^{N+\mu}z)(t) \text{ a.e. in } [s,t^*],$$

if $V_s^{N,\mu} \in \mathscr{L}(L^2(s,t^*;\mathbb{R}^m), L^2(s,t^*;\mathbb{R}^m))$ and $W_s^{N,\mu} \in \mathscr{L}(Z,L^2(s,t^*;\mathbb{R}^m))$ are defined analogously to $V_{t_0}^{N,\mu}$ and $W_{t_0}^{N,\mu}$ in (2.11) and (2.12) respectively. For $z \in Z$ the optimal trajectories $S^{N,\mu}(t,s)z$ corresponding to $J(s,P^{N+\mu}z,u)$ are then given by

$$(2.14)\qquad S^{N,\mu}(t,s)P^{N+\mu}z = T^{N+\mu}(t,s)P^{N+\mu}z - \int_s^t T^{N+\mu}(t,\eta)\mathscr{B}^N(\eta)((V_s^{N,\mu})^{-1}W^{N,\mu}P^{N+\mu}z)(\eta)d\eta.$$

In [6] it is verified that $S^{N,\mu}(t,s)$ is an evolution operator on Δ for each N,μ and moreover that $\tilde{u}^{N,\mu}$ is also given by

$$(2.15)\qquad \tilde{u}^{N,\mu}(t) = -C^{-1}(t)(\mathscr{B}^N(t))^*\Pi^{N,\mu}(t)S^{N,\mu}(t,t_0)P^{N+\mu}z, \text{ a.e.,}$$

with

$$(2.16)\qquad \Pi^{N,\mu}(t)P^{N+\mu}z = (T^{N+\mu})^*(t^*,t)F^{N+\mu}S^{N,\mu}(t^*,t)P^{N+\mu}z + \int_t^{t^*}(T^{N+\mu})^*(\eta,t)\mathscr{D}^{N+\mu}(\eta)S^{N,\mu}(\eta,t)P^{N+\mu}z\,d\eta,$$

$$\text{for } t_0 \le t \le t^*,\ z \in Z.$$

The basis for the numerical approximation scheme will be (2.10) and (2.14) together with the following Riccati integral equation for $\Pi^{N,\mu}$.

$$(2.17)\quad \Pi^{N,\mu}(t)P^{N+\mu}z = (T^{N+\mu})^*(t^*,t)F^{N+\mu}(t^*,t)P^{N+\mu}z + \int_t^{t^*}(T^{N+\mu})^*(\eta,t)[\mathscr{D}^{N+\mu}(\eta) - \Pi^{N,\mu}(\eta)\mathscr{B}^N(\eta)C^{-1}(\eta)(\mathscr{B}^N(\eta))^*\Pi^{N,\mu}(\eta)]T^{N+\mu}(\eta,t)P^{N+\mu}z\,d\eta.$$

Since Z is separable, so that $T^*(\cdot,\cdot)$ is strongly measurable, (2.17) is also a direct consequence of the results in [6]; moreover

$$\Pi^{N,\mu}(t) \text{ is nonnegative and selfadjoint.}$$

To establish the approximation results we adopt the following conventions:

(i) 0 as a superscript may be dropped,

(ii) $b = \sup_{t\in[t_0,t^*]} ||B(t)||$,

(iii) $c = \sup_{t\in[t_0,t^*]} ||C(t)||$,

(iv) $d = \sup_{t\in[t_0,t^*]} ||D(t)||$,

(v) $f = ||F||_{\mathscr{L}(Z,Z)}$,

(vi) we assume that $M \leq \overline{M}$, $\omega \leq \overline{\omega}$,

(vii) by (H1)(i), (iii) there exists a real-valued function ρ such that

$$|T(t,s)z - T^N(t,s)P^N z| \leq \rho(N,z), \text{ uniformly in } \Delta;$$

indeed $\rho(N,z) = \bar{\rho}(N,z) + \bar{M}e^{\bar{\omega}(t^*-t_0)}|P^N z - z|$,

(viii) by (H1)(iii) there exists a real valued function $\tilde{\rho}$ such that

$$||T(t,s)Q^N - T^{N+\mu}(t,s)Q^N||_{\mathscr{L}(\mathbb{R}^n,Z)} \leq \tilde{\rho}(\mu+N,Q^N),$$

indeed one can let

$$\tilde{\rho}(\mu+N,Q^N) = \sqrt{n} \max_{j=1,\ldots,n} \bar{\rho}(N+\mu,((Q_0^N)_j,(Q_1^N)_j)),$$

(ix) the constants K_i to be used below depend on the following variables of $(\mathscr{P}^{N,\mu})$

$$K_i = K_i(n,A_i,f,d,b,c,q,t_0,t^*)$$

and are calculated explicitly in the proofs.

<u>Lemma 2.4</u>.

(a) $||W_{t_0}^{N,\mu}||_{\mathscr{L}(Z;C(t_0,t^*;\mathbb{R}^m))} \leq bq\bar{M}^2 e^{2\bar{\omega}(t^*-t_0)}(d(t^*-t_0) + f) \overset{\text{def}}{=} K_1$.

(b) $||(V_{t_0}^{N,\mu})^{-1}||_{\mathscr{L}(L^2(t_0,t^*;\mathbb{R}^m),L^2(t_0,t^*;\mathbb{R}^m))} \leq c^{-1}$.

(c) There exist constants K_2 and K_3 such that for all $z \in Z$

$$\sup_{t\in[t_0,t^*]} |(W_{t_0} z)(t) - (W_{t_0}^{N,\mu} P^{N+\mu} z)(t)|_{\mathbb{R}^m} \leq$$

$$\leq \rho_Q(N)|z|k_1 + \tilde{\rho}(N+\mu,Q^N)|z|K_2 + \rho(N+\mu,z)K_3$$

$$+ ||P^{N+\mu}Q_0 - Q_0||\ |z|k_2,$$

where

$$k_1 = b\sqrt{n}\ \bar{M}^2 e^{2\bar{\omega}(t^*-t_0)}[(2\bar{\omega})^{-1}d+f]$$

$$k_2 = bq\sqrt{n}\ \bar{M}^2 e^{2\bar{\omega}(t^*-t_0)}[d(t^*-t_0)(2\bar{\omega})^{-1} + f].$$

(d) There exists a constant K_4 such that for all $w \in L^2(t_0,t^*;\mathbb{R}^m)$,

$$\sup_{t\in[t_0,t^*]} |((V_{t_0} - V_{t_0}^{N,\mu})w)(t)| \leq \rho_Q(N)k_3|w|_{L^2} + \tilde{\rho}(N+\mu,Q^N)K_4|w|_{L^2}$$

$$+ ||Q_0 - P^{N+\mu}Q_0||\ k_4|w|_{L^2},$$

where

$$k_3 = b^2\sqrt{n}\ \bar{M}^2(1+q)e^{2\bar{\omega}(t^*-t_0)}\left(\frac{f}{\sqrt{2\bar{\omega}}} + \frac{d}{\sqrt{\bar{\omega}^3}}\right),$$

and

$$k_4 = b^2\sqrt{n}\ \bar{M}\ q^2 e^{\bar{\omega}(t^*-t_0)}\left[f\left(\frac{1}{\sqrt{2\bar{\omega}}} + (t^*-t_0)^{1/2}\right) + d\left(\frac{1}{\sqrt{2\bar{\omega}^3}} + \frac{(t^*-t_0)^{1/2}}{\bar{\omega}}\right)\right].$$

Remark 2.2. It is simple to check that the same estimates as in the previous lemma also hold for $V_s^{N,\mu} \in \mathcal{L}(L^2(s,t^*;\mathbb{R}^m), L^2(s,t^*;\mathbb{R}^m))$ and $W_s^{N,\mu} \in \mathcal{L}(Z,L^2(s,t^*;\mathbb{R}^m))$, for $s \in [t_0,t^*]$.

Theorem 2.1. For the optimal controls $\tilde{u}^{N,\mu}$ and $\tilde{u}$ we have the following L^2-convergence result:

$$|\tilde{u}-\tilde{u}^{N,\mu}|_{L^2(t_0,t^*;\mathbb{R}^m)} \leq c^{-1}(t^*-t_0)^{1/2}[\rho_Q(N)|z|k_1 + \tilde{\rho}(N+\mu,Q^N)|z|K_2$$
$$\rho(N+\mu,z)K_3 + ||P^{N+\mu}Q_0-Q_0||\ |z|k_2] + c^{-2}(t^*-t_0)K_1|z|[\rho_Q(N)k_3$$
$$+ \tilde{\rho}(N+\mu,Q^N)K_4 + ||Q_0-P^{N+\mu}Q_0||k_4].$$

Proof. The proof, using Lemma 2.4, follows from the following simple estimate

$$|\tilde{u}-\tilde{u}^{N,\mu}|_{L^2} \leq |V_{t_0}^{-1}W_{t_0}z - (V_{t_0}^{N,\mu})^{-1}W_{t_0}^{N,\mu}P^{N+\mu}z|_{L^2} =$$
$$= |V_{t_0}^{-1}W_{t_0}z - V_{t_0}^{-1}W_{t_0}^{N,\mu}P^{N+\mu}z|_{L^2} + |V_{t_0}^{-1}W_{t_0}^{N,\mu}P^{N+\mu}z - (V_{t_0}^{N,\mu})^{-1}W_{t_0}^{N,\mu}P^{N+\mu}z|_{L^2} \leq$$
$$\leq c^{-1}|W_{t_0}z - W_{t_0}^{N,\mu}P^{N+\mu}z|_{L^2} + |(V_{t_0}^{N,\mu})^{-1}(V_{t_0}-V_{t_0}^{N,\mu})(V_{t_0})^{-1}W_{t_0}^{N,\mu}P^{N+\mu}z|_{L^2} \leq$$
$$\leq c^{-1}|W_{t_0}z - W_{t_0}^{N,\mu}P^{N+\mu}z|_{L^2} +$$
$$+ c^{-2}||V_{t_0} - V_{t_0}^{N,\mu}||_{\mathcal{L}(L^2(t_0,t^*;\mathbb{R}^m),L^2(t_0,t^*;\mathbb{R}^m))}K_1(t^*-t_0)^{1/2}|z|.$$

Remark 2.4. Although it is not difficult to show that the controls $\tilde{u}^{N,\mu}$; $N,\mu = 0,1,\ldots$ are continuous, if $B(\cdot)$ and $C(\cdot)$ are and that $V^{N,\mu}$ is invertible in $\mathcal{L}(C(t_0,t^*;\mathbb{R}^m), C(t_0,t^*;\mathbb{R}^m))$ it does not seem possible to find a uniform bound on $\|(V^{N,\mu})^{-1}\|_{\mathscr{L}(C(t_0,t^*;\mathbb{R}^m),C(t_0,t^*;\mathbb{R}^m))}$. We shall, however, consider the question of uniform convergence of the controls in Theorem 2.4.

Remark 2.4. The use of Theorem 2.1 will be demonstrated for the case where the subspaces Z^N are chosen as subspaces of spline functions, for example. Then, if one is merely interested in convergence of the optimal controls $\tilde{u}^{N,\mu}$ one may put $\mu = 0$ and Theorem 2.1 will guarantee L^2-convergence of $\tilde{u}^{N,0}$ to $\tilde{u}$. However, if the initial data $z \in Z$ are picked sufficiently smooth one would expect to find higher order estimates on the rate of convergence. Unfortunately even for smooth initial data, one still has to deal with the "jump" operator Q_0 used in the variation-of-constants formula (compare (2.4) and $(\mathscr{P}^{N,\mu})$): given any $\varepsilon > 0$ one can determine $N > 0$, such that for all $\mu = 1,2,\ldots$

$$\|Q_0-Q^N\|[c^{-1}(t^*-t_0)^{1/2}|z|k_1 + c^{-2}(t^*-t_0)K_1|z|k_3] + \|P^{N+\mu}Q_0-Q_0\|[c^{-1}(t^*-t_0)^{1/2}|z|k_2 + c^{-2}(t^*-t_0)K_1|z|k_4] < \varepsilon.$$

Fixing N, Theorem 2.1 guarantees that the optimal controls converge at a rate given by $\tilde{\rho}(N+\mu,Q^N)$ and $\rho(N+\mu,z)$ into the ε-bound as $\mu \to \infty$.

Remark 2.5. It can be seen easily that Theorem 2.1 remains true if $\tilde{u}^{N,\mu}$ and $\tilde{u}$ are replaced by $\tilde{u}_s^{N,\mu}$ and $\tilde{u}_s$ as defined in (2.13).

For the optimal trajectories $S(t,s)z$, with $(t,s) \in \Delta$, we have the following estimate.

Theorem 2.2.

(a) There exists a constant K_5, independent of N and μ, such that

$$|S^{N,\mu}(t,s)P^{N+\mu}z| \leq K_5|z| \quad \text{for all } (t,s) \in \Delta \text{ and } \mu,\ N = 0,1,2,\ldots .$$

(b) There exists a constant K_6 such that

$$|S(t,s)z - S^{N,\mu}(t,s)P^{N+\mu}z| \leq \rho(N+\mu,z) + \tilde{\rho}(N+\mu,Q^N)|z|K_6 + |\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}k_5 + |z|\rho_Q(N)k_6,$$

where

$$k_5 = b\overline{M}\left(\frac{1}{2\overline{\omega}}\left(e^{2\overline{\omega}(t^*-t_0)}-1\right)\right)^{1/2}$$

$$k_6 = bc^{-1}K_1(t^*-t_0)^{1/2}\overline{M}\left(\frac{1}{2\overline{\omega}}\left(e^{2\overline{\omega}(t^*-t_0)}-1\right)\right)^{1/2}.$$

Theorem 2.3. For all $z \in Z$ and $t \in [t_0,t]$

(a) $||\Pi^{N,\mu}(t)P^{N+\mu}z|| \leq \overline{M}e^{\overline{\omega}(t^*-t_0)}K_5|z| \quad (f+d(t^*-t_0))$.

(b) $$\langle\Pi(t)z - \Pi^{N,\mu}(t)z^{N+\mu},y\rangle \leq (d(t^*-t_0) + f)\{\overline{\rho}(N+\mu,y)|z|K_5 + \overline{M}e^{\overline{\omega}(t^*-t_0)}|y|[\rho(N+\mu,z) + |z|\tilde{\rho}(N+\mu,Q^N)K_6 + |\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}k_5 + |z|\rho_Q(N)k_6 + |z|K_5||Q_0-P^{N+\mu}Q_0||]\},$$

for all $y \in Z$.

For $\mu = 0$ we have the following corollary to Theorem 2.1-2.3.

Corollary 2.1. There exist constants K_7, K_8, K_9 such that

(a) $|\tilde{u}-\tilde{u}^{N,0}|_{L^2(t_0,t^*;\mathbb{R}^m)} \leq K_7[\tilde{\rho}(N,Q_0)|z| + \rho(N,z) + \rho_Q(N)|z|]$,

(b) $|S(t,s)z - S^{N,0}(t,s)P^Nz| \leq K_8[\tilde{\rho}(N,Q_0)|z| + \rho(N,z) + \rho_Q(N)|z|]$,

(c) $\langle \Pi(t)z - \Pi^{N,0}P^Nz, y\rangle \leq K_9[\bar{\rho}(N,y)|z| + \rho(N,z)|y|$

$+ |y||z|(\rho_Q(N) + \tilde{\rho}(N,Q_0))]$.

Proof. By Theorem 2.1 we have

$$(2.18)\quad |\tilde{u}-\tilde{u}^{N,0}|_{L^2(t_0,t^*;\mathbb{R}^m)} \leq \tilde{K}_7[\rho_Q(N)|z| + \tilde{\rho}(N,Q^N)|z| + \rho(N,z) + \|P^NQ_0 - Q_0\|\ |z|].$$

Since $Q^N\mathbb{R}^n \subset Z^N$ and since P^N is an orthogonal projection, $|P^NQ_0\eta - Q_0\eta| < ||Q^N - Q_0||\ |\eta|$ for all η, which implies that

$$(2.19)\qquad ||P^NQ_0 - Q_0|| < \rho_Q(N).$$

Also, for $(t,s) \in \Delta$ we find

$$(2.20)\quad \tilde{\rho}(N,Q^N) \leq ||T(t,s)(Q^N-Q_0)|| + ||(T(t,s) - T^N(t,s))Q_0|| +$$

$$+ ||T^N(t,s)(Q_0-Q^N)|| \leq 2\bar{M}e^{\bar{\omega}(t-s)}\rho_Q(N) + \tilde{\rho}(N,Q_0).$$

Estimates (2.18)-(2.20) imply (a). Similar calculations prove (b) and (c).

<u>Corollary 2.2</u>. For the payoff J the following estimate holds:

$$|J(t_0,P^{N+\mu}z,\tilde{u}^{N,\mu}) - J(t_0,z,\tilde{u})| \leq$$

$$\leq [(t^*-t_0)d + f]K_5|z|\{2[\rho(N+\mu,z) + \tilde{\rho}(N+\mu,Q^N)|z|K_6 + |\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}k_5 + |z|\rho_Q(N)k_6] +$$

$$+ ||P^{N+\mu}Q_0 - Q_0||K_5|z|\} + 2|\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}K_1c^{-1}(t^*-t_0)^{1/2}\sup||C(t)|||z|.$$

Finally, we discuss the convergence of $\tilde{u}^{N,\mu}$ to $\tilde{u}$ in the supremum-norm. For the sake of a simpler representation we restrict our attention to the case $\mu = 0$.

<u>Theorem 2.4</u>. If for all $z \in Z$, $\lim_{N\to\infty} \bar{\rho}(N,z) = 0$ and if $B(t)$ and $C(t)$ are continuous in t then

$$\lim_{N\to\infty} \sup_{t\in[t_0,t^*]} |\tilde{u}(t) - \tilde{u}^{N,0}(t)| = 0.$$

We draw the readers attention to the fact that all convergence results were obtained avoiding any specific information about the adjoint evolution operator or its generator. This is quite important, since the properties of the adjoint evolution operator are unfavorable to constructing approximation schemes: if $\mathcal{A}^*(t)$ denotes the infinitesimal generator of the adjoint evolution operator, then

$$\bigcap_{t\geq t_0} \mathrm{Dom}\,(\mathcal{A}^*(t)) \quad \text{need not be dense in} \quad Z$$

(see [4]) and for autonomous (FDE) $\mathrm{Dom}(\mathcal{A}^*)$ consists of all elements

$(\eta,\phi) \in Z$ with ϕ absolutely continuous on $[-r_i, r_{i-1})$, for $i = 1,\ldots,\ell$, and with jumps at r_i determined by A_i^*. It would therefore be quite difficult to find a sequence of operators $T_*^N(t,s)$ satisfying properties analogous to (H1) with $T^N(t,s)$ and $T(t,s)$ replaced by $T_*^N(t,s)$ and $T(t,s)^*$ respectively; indeed we shall see shortly that $Z^N \subset \mathrm{Dom}(\mathscr{A}(t))$ is a very convenient property for showing that $T^N(t,s)$ converges to $T(t,s)$, but the analogous hypothesis $Z^N \subset \mathrm{Dom}(\mathscr{A}^*(t))$ would rarely be satisfied.

3. Spline Approximation Schemes

In this section we apply the results of the previous one to subspaces of spline functions. There are three subsections: (α) Generalities, (β) Linear-spline-functions, (γ) Cubic-spline functions.

(α) Generalities. Spline approximations for (FDE) have been developed in [3] and we shall use these results here. Throughout, we assume (2.3) to be autonomous, so that A_i, $i = -1,0,\ldots$, are independent of t. We recall that in this case the solution evolution operator becomes a semigroup via $T(t,s)z = T(t-s)z$ for $z \in Z$ and $(t,s) \in \Delta$, whose infinitesimal generator $\mathscr{A}$ is given by $\mathscr{A}(\phi(0),\phi) = (L(\phi),\dot{\phi})$, where $\mathrm{Dom}(\mathscr{A}) = \{(\eta,\phi) \mid \phi \in W^{1,2}(-r,0;\mathbb{R}^n),\ \phi(0) = \eta\}$. We specify a weighting function for the norm of Z by

$$g(s) = j \quad \text{for} \quad s \in [-r_{\ell-j+1}, -r_{\ell-j}), \quad \text{for} \quad j = 1,\ldots,\ell.$$

Obviously Z_1 and Z_g are equivalent Banach spaces, since $\ell^{-1/2}|(\eta,\phi)|_{Z_g} \le |(\eta,\phi)_{Z_1}| \le |(\eta,\phi)|_{Z_g}$. We continue to drop the subscript g if only the set-theoretic or topological structures of Z are important. Next, we repeat a general result from [3], and call $\{Z^N, P_g^N, \mathscr{A}^N\}$, $N = 1,2,\ldots$ an approximation scheme if $\{Z^N\}$ is a sequence of closed linear subspaces of Z_g, $\{P_g^N\}$ is the sequence of orthogonal projections, $P_g^N\colon Z_g \to Z^N$ and $\{A^N\}$ is a sequence of operators $Z_g \to Z^N$.

<u>Theorem 3.1</u>. Let $\{Z^N, P_g^N, \mathscr{A}^N\}$ be an approximation scheme satisfying

(i) $Z^N \subset \mathrm{Dom}(\mathscr{A})$, $N = 1,2,\ldots$

(ii) $\mathscr{A}^N = P_g^N \mathscr{A} P_g^N$, $N = 1,2,\ldots$

(iii) (a) $\lim_{N\to\infty} P_g^N z = z$ in Z for all $z \in Z$,

(b) for some integer $k \geq 1$ we have

$\lim_{N\to\infty} L(\psi^N) = L(\psi)$ in $\mathbb{R}^n$ and

$\lim_{N\to\infty} (\psi^N)' = \psi'$ in L^2 for all $\psi \in C^k$, where ψ^N is defined by $P_g^N \hat{\psi} = (\psi^N(0), \psi^N)$.

Then each $\mathscr{A}^N$ is the infinitesimal generator of a C_0-semigroup $T^N(t)$, $t \geq 0$, such that

$$T^N(t) Z^N \subset Z^N, \quad N = 1,2,\ldots, \quad t \geq 0$$

$$\lim_{N\to\infty} T^N(t) z = T(t) z, \quad \text{in } Z$$

and

$$\|T^N(t)\|_{Z_g} \leq e^{\tilde{\omega} t}, \quad \|T(t)\|_{Z_g} \leq e^{\tilde{\omega} t}$$

where

$$\tilde{\omega} = \frac{\ell+1}{2} + \|A_0\| + \frac{1}{2} \sum_{i=1}^{\ell} \|A_i\|^2 + \frac{1}{2} \int_{-r}^{0} \|A(s)\|^2 ds.$$

We carefully avoided $\mathscr{A}^*$, the adjoint of $\mathscr{A}$; however we shall need

<u>Lemma 3.1</u>. The infinitesimal generators $\mathscr{A}_N^*$ of $(T^N(t))^*$, the adjoint semigroup of $T^N(t)$ generated by $P_g^N \mathscr{A} P_g^N$, are given by

$$\mathrm{Dom}(\mathscr{A}_N^*) = Z$$

and

$$\mathscr{A}_N^* = (P_g^N \mathscr{A} P_g^N)^*.$$

<u>Proof</u>. Since $\mathscr{A} P_g^N$ is closed and defined on Z it is bounded; therefore $P_g^N \mathscr{A} P_g^N$ is bounded and so is $(P_g^N \mathscr{A} P_g^N)^*$ and $\mathrm{Dom}(P_g^N \mathscr{A} P_g^N)^* = Z$. The second claim follows from general semigroup theory [10, pp. 277].

We need one more condition on the subspaces Z^N:

$$\text{(H4)} \quad \begin{cases} \dim Z^N = k_N < \infty \quad \text{and} \\ \text{for each } N = 1,2,\ldots \text{ there exists an integer } \mu > 0 \text{ such that} \\ \qquad Z^N \subset Z^{N+\mu}. \end{cases}$$

We now turn to a discussion of the variation of constants formula (2.9), the feedback laws (2.13) and (2.15) and the Riccati integral equation (2.17); the assumptions used in the rest of this section are (H3), (H4) and the assumptions of Theorem 3.1.

The fact that by (H3) the columns of Q^N are in Z^N, together with (H4) and $T^N(t)Z^N \subset Z^N$ imply that for each N there exists an integer μ such that the right-hand-side of (2.9) is in the finite dimensional subspace $Z^{N+\mu}$ and therefore,

$$
(3.1)\quad
\begin{aligned}
&\dot{z}^{N,\mu}(t) = \mathscr{A}^{N+\mu} z^{N,\mu}(t) + Q^N B(t)u(t), \quad t_0 \le t \le t^* \\
&z^{N,\mu}(t_0) = P_g^{N+\mu} z .
\end{aligned}
$$

By a similar argument we find that $\Pi^{N,\mu}(\cdot)$ satisfies in $Z^{N+\mu}$ the Riccati differential equation

$$
(3.2)\quad
\begin{cases}
\dfrac{d\Pi^{N,\mu}(t)}{dt} = (\mathscr{A}^{N+\mu})^* \Pi^{N,\mu}(t) + \Pi^{N,\mu}(t)\mathscr{A}^{N+\mu} - \\
\quad - [\mathscr{D}^{N+\mu}(t) - \Pi^{N,\mu}(t)\mathscr{B}^N(t)C^{-1}(t)(\mathscr{B}^N)^*(t)\,\Pi^{N,\mu}(t)] \\
\qquad\qquad\qquad \text{for } t_0 \le t \le t^* , \\
\Pi^{N,\mu}(t^*) = F^{N+\mu}.
\end{cases}
$$

We also recall the feedback law

$$
(3.3)\qquad \tilde{u}^{N,\mu}(t) = -C^{-1}(t)(\mathscr{B}^N)^*(t)\Pi^{N,\mu}(t)S^{N,\mu}(t,t_0)P_g^{N+\mu}z,
$$

where $S^{N,\mu}(\cdot,t_0)P_g^{N+\mu}z$ is the optimal trajectory corresponding to $(\mathscr{P}^{N,\mu})$.

To approximate $(\mathscr{P})$ by the finite dimensional problems $(\mathscr{P}^{N,\mu})$ we yet have to express the various operators in (3.1)-(3.3) with respect to some bases in Z^N.

<u>Remark 3.1</u>. For the reader, who cares to follow the calculations carried out in this subsection, or the potential applicant of the resulting finite dimensional linear quadratic control problem, it might be helpful to think of $\mathbb{R}^n$-vectors as $n \times n$ diagonal-valued matrices.

For each $N = 1,2,\ldots$ we now choose a basis $(\hat{\beta}_1^N,\ldots,\hat{\beta}_{k_N}^N)$ of Z^N. From $Z^N \subset \mathrm{Dom}(\mathscr{A})$ it follows that $\hat{\beta}_i^N = (\beta_i^N(0),\beta_i^N)$ for $i = 1,\ldots,k_N$, with $\beta_i^N \in W^{1,2}$. We shall need the matrix functions

$$\beta^N = (\beta_1^N,\ldots,\beta_{k_N}^N)$$

and

$$\hat{\beta}^N = (\hat{\beta}_1^N,\ldots,\hat{\beta}_{k_N}).$$

Each element $z^N \in Z^N$ can be expressed as

$$z^N = \hat{\beta}^N\alpha^N, \text{ for } \alpha^N = \mathrm{col}(\alpha_1^N,\ldots,\alpha_{k_N}^N) \in \mathbb{R}^{k_N}$$

or in terms of elements as

$$z^N = \left(\sum_{i=1}^{k_N} \beta_i^N(0)\alpha_i^N, \quad \sum_{i=1}^{k_N} \beta_i^N\alpha_i^N\right).$$

The matrix representation of $\mathscr{A}^N: Z^N \to Z^N$, denoted by A^N and the coordinate vector of $P_g^N z$, for $z = (\eta,\phi) \in Z$ have been calculated in [3]. To present this result, which is a simple consequence of $P_g^N(\eta,\phi) - (\eta,\phi) \perp Z_g^N$, we define the matrices

$$J^N = \langle\hat{\beta}^N,\hat{\beta}^N\rangle_{Z_g} \overset{\mathrm{def}}{=} \beta^N(0)^*\beta^N(0) + \int_{-r}^0 \beta^N(s)^*\beta^N(s)g(s)ds$$

$$h^N(\eta,\phi) = \langle\hat{\beta}^N,(\eta,\phi)\rangle_{Z_g} \overset{\mathrm{def}}{=} \beta^N(0)^*\eta + \int_{-r}^0 \beta^N(s)^*\phi(s)g(s)ds$$

and

$$(3.4)\quad H^N = h^N(L(\beta^N),\dot{\beta}^N) = \beta^N(0)^* L(\beta^N) + \int_{-r}^{0} \beta^N(s)^* \dot{\beta}^N(s)g(s)ds.$$

Then, if $P_g^N(\eta,\phi) = \hat{\beta}^N\alpha^N$, the coordinate vector α^N is given by

$$(3.5)\qquad \alpha^N = (J^N)^{-1}h^N(\eta,\phi)$$

and the matrix representation of $\mathscr{A}^N$ by

$$(3.6)\qquad A^N = (J^N)^{-1}H^N.$$

If for any N, one choses μ satisfying (H4), then the columns of Q^N are in $Z^{N+\mu}$ and there exists a vector $\delta^{N,\mu} \in \mathbb{R}^{k_{N+\mu}}$ such that

$$(3.7)\qquad Q^N = \hat{\beta}^{N+\mu}\delta^{N,\mu}.$$

Since we think of the approximation in N as chosen by the user according to some desired accuracy which can be achieved by fixing N sufficiently large and then by letting $\mu \to \infty$, the following formula will be useful

$$(3.8)\qquad \delta^{N,\mu} = (J^{N+\mu})^{-1}\langle\hat{\beta}^{N+\mu},\hat{\beta}^N\rangle_{Z_g}\delta^{N,0}.$$

Next, we turn to $(Q^{N+\mu})^*\colon Z^{N+\mu} \to \mathbb{R}^n$. For $\hat{\phi}^{N+\mu} = (\phi^{N+\mu}(0),\phi^{N+\mu}) \in Z^{N+\mu}$ define $\gamma^{N+\mu} \in \mathbb{R}^{k_{N+\mu}}$, and $\eta \in \mathbb{R}^n$ by

$$\hat{\phi}^{N+\mu} = \hat{\beta}^{N+\mu}\gamma^{N+\mu} \quad \text{and} \quad (Q^{N+\mu})^*\hat{\phi}^{N+\mu} = \eta.$$

Then for all $x \in \mathbb{R}^n$

$$\langle Q^{N+\mu}x, \hat{\phi}^{N+\mu}\rangle_{Z_g} = (x, (Q^{N+\mu})^*\hat{\phi}^{N+\mu})$$

or, equivalently

$$\langle \hat{\beta}^{N+\mu}\delta^{N,\mu}x,\ \hat{\beta}^{N+\mu}\gamma^{N+\mu}\rangle_{Z_g} = (x, \eta).$$

We use Remark 3.1 and easily deduce

$$(\delta^{N,\mu})^*\ J^{N+\mu}\gamma^{N+\mu} = \eta.$$

Therefore, the matrix representation $[(Q^N)^*]$ of Q^{N*} is given by

(3.9) $$[(Q^N)^*] = (\delta^{N,\mu})^*J^{N+\mu} = (\delta^{N,0})^*\langle\hat{\beta}^N, \hat{\beta}^{N+\mu}\rangle_{Z_g}.$$

Let A_*^N denote the matrix representation of $(P_g^N \mathcal{A} P_g^N)^*$, then

(3.10) $$A_*^N = (J^N)^{-1}(A^N)^*J^N.$$

Indeed, let $\hat{\phi} = (\phi(0), \phi) \in Z^N$ and $\hat{\psi} = (\psi(0), \psi) \in Z^N$ be given by

$$\hat{\phi} = \hat{\beta}^N\rho^N \quad \text{and} \quad \hat{\psi} = \hat{\beta}^N\sigma^N, \quad \text{for } \rho^N \in \mathbb{R}^{k_N}, \sigma^N \in \mathbb{R}^{k_N}.$$

Then the equality

$$\langle (P_g^N \mathscr{A} P_g^N)^* \hat{\phi}, \hat{\psi} \rangle_{Z_g} = \langle \hat{\phi}, P_g^N \mathscr{A} P_g^N \hat{\psi} \rangle_{Z_g}$$

implies

$$\langle \hat{\beta}^N A_*^N \rho^N, \hat{\beta}^N \sigma^N \rangle_{Z_g} = \langle \hat{\beta}^N \rho^N, \hat{\beta}^N A^N \sigma^N \rangle_{Z_g},$$

so that

$$\left\langle \sum_{i,j=1}^{k_N} \hat{\beta}_i^N (A_*^N)_{i,j} \rho_j^N, \sum_{v=1}^{k_N} \hat{\beta}_v^N \sigma_v^N \right\rangle_{Z_g} =$$

$$= \left\langle \sum_{v=1}^{k_N} \hat{\beta}_v^N \rho_v^N, \sum_{i,j}^{k_N} \hat{\beta}_i^N (A^N)_{i,j} \sigma_j^N \right\rangle_{Z_g},$$

which upon interchanging v by j on the right hand side becomes

$$\sum_{i,j,v}^{k_N} \langle \hat{\beta}_i^N, \hat{\beta}_v^N \rangle_{Z_g} (A_*^N)_{i,j} \rho_j^N \sigma_v^N = \sum_{i,j,v}^{k_N} \langle \hat{\beta}_j^N, \hat{\beta}_i^N \rangle_{Z_g} (A^N)_{i,v} \rho_j^N \sigma_v^N .$$

Since $\hat{\phi}$ and $\hat{\psi}$ are arbitrary, the last equation has to hold for all vectors ρ^N and σ^N, and therefore

$$\sum_{i=1}^{k_N} \langle \hat{\beta}_i^N, \hat{\beta}_v^N \rangle_{Z_g} (A_*^N)_{i,j} = \sum_{i=1}^{k_N} \langle \hat{\beta}_j^N, \hat{\beta}_i^N \rangle (A^N)_{i,v} \quad \text{for all } j \text{ and } v = 1,\dots,k_N.$$

This equality implies (3.10).

We notice, of course, that if $\hat{\beta}_1,\dots,\hat{\beta}_{k_N}$ were an orthonormal basis then A_*^N would equal $(A^N)^*$. To find the matrix representation $[D^N(t)]$ of $\mathscr{D}^N(t) = P_g^N \mathscr{D}(t) P_g^N$ we let $\hat{\phi}^N = (\phi^N(0), \phi^N) \in Z^N$ and define $\alpha^N \in \mathbb{R}^{k_N}$ and $\gamma^N \in \mathbb{R}^{k_N}$ by

$$\hat{\phi}^N = \hat{\beta}^N \alpha^N \quad \text{and} \quad \mathscr{D}^N(t)\hat{\phi}^N = \hat{\beta}^N \gamma^N.$$

Therefore

$$\mathscr{D}^N(t)\hat{\phi}^N = P_g^N \mathscr{D}(t)\hat{\phi}^N = P_g^N(D(t)\phi^N(0),0)$$

and by (3.5)

$$\gamma^N = (J^N)^{-1} h^N(D(t)\phi^N(0),0).$$

But

$$h^N(D(t)\phi^N(0),0) = h^N(D(t)\beta^N(0)\alpha^N,0) = \tilde{D}^N(t)\alpha^N$$

where

$$\tilde{D}^N(t) = \beta^N(0)^* D(t)\beta^N(0),$$

so that

$$\gamma^N = (J^N)^{-1}\tilde{D}^N(t)\alpha^N$$

or

$$[D^N(t)] = (J^N)^{-1}\tilde{D}^N(t). \tag{3.11}$$

In a similar manner we see that the representation $[F^N]$ of $F^N = P_g^N \mathscr{F} P_g^N$ is given by

$$[F^N] = (J^N)^{-1}\tilde{F}^N,$$

where

$$\tilde{F}^N = \beta^N(0)^* F \beta^N(0).$$

To write equations (3.1)-(3.3) in terms of the coordinate representative with respect to the basis chosen, we let $\pi^{N,\mu}(t)$ denote the matrix representation of the operators $\Pi^{N,\mu}(t)\colon Z^N \to Z^N$ and let $w^{N,\mu}(t) = w^{N,\mu}(t;u)$ and $w_0^{N,\mu}$ be defined by

$$z^{N,\mu}(t) = \hat{\beta}^{N+\mu} w^{N,\mu}(t) \quad \text{and} \quad P_g^{N+\mu} z = \hat{\beta}^{N+\mu} w_0^{N,\mu}.$$

Then (3.1)-(3.3) are equivalent to

$$(3.12)\quad \begin{cases} \dot{w}^{N,\mu}(t) = A^{N+\mu} w^{N,\mu}(t) + \delta^{N,\mu} B(t) u^{N,\mu}(t) \quad \text{for } t_0 \le t \le t^* \\ w^{N,\mu}(t_0) = w_0^{N,} \\ \frac{d}{dt}\pi^{N,\mu}(t) = A^*_{N+\mu}\pi^{N,\mu}(t) + \pi^{N,\mu}(t) A^{N+\mu} - \\ \qquad - [D^{N+\mu}(t)] - \pi^{N,\mu}(t)\delta^{N,\mu} B(t) C(t) B^*(t) \delta^{N,\mu^*} J^{N+\mu}\pi^{N,\mu}(t) \\ \pi^{N,\mu}(t^*) = [F^{N+\mu}] \\ u^{N,\mu}(t) = -C^{-1}(t) B^*(t) \delta^{N,\mu^*} J^{N+\mu}\pi^{N,\mu}(t) w^{N,\mu}(t). \end{cases}$$

We close this subsection with a final remark on the choice of the operator Q^N.

<u>Remark</u>. The natural possibilities of choosing Q^N are to

(α) either take $Q^N = P^N_g Q_0$, which by (3.5) implies that

$$\delta^{N,0} = (J^N)^{-1}\beta^N(0)^T$$

(β) or, if the subspaces are chosen as spline functions, to take the representation of $Q^N: \mathbb{R}^n \to Z$ as the interpolating spline, which at the knot $t = 0$ takes the value I (identity matrix) and 0 (zero-matrix) on the other knots.

The choice between (α) and (β) has to be made on the ground of getting the best convergence for $\tilde{\rho}(N,Q^N)$. Condition (H3) is checked in essentially the same manner for (α) and (β).

(β) <u>Linear spline functions</u>.

We begin this subsection with a brief discussion on the rate of convergence of the approximating semigroups $T^N(t)$ constructed in Theorem 3.1. In [3] it is shown that general semigroup theory provides the following estimate:

There exists a constant $\tilde{M} = \tilde{M}(t^*,A_i,\lambda_0)$ such that

$$(3.13)\quad |(T^N(t)-T(t))z| \le M\{|(\mathscr{A}^N-\mathscr{A})z| + \int_0^{t^*} |[\mathscr{A}^N-\mathscr{A}]T(s)(\lambda_0 I-\mathscr{A})z|ds + |[\mathscr{A}^N-\mathscr{A}]T(t)z|\},$$

for all $t \in [0,t^*]$ and $z \in \mathrm{Dom}(\mathscr{A}^2)$.

To give estimates on the rate of convergence for our problem we use (3.13) together with results from the theory of spline functions to estimate $\mathscr{A}^N \to \mathscr{A}$. In spline analysis the estimate on the rate of convergence of an interpolating spline always contains higher order derivatives of the function that it interpolates. This constitutes an essential problem for choosing Q^N and estimating $\tilde{\rho}(N,Q^N)$, since (H3) does not allow to pick the representative of Q^N arbitrarily smooth.

Although estimate (3.13) (which was derived from general semigroup-theory) might be too weak for the special case of (FDE), it nevertheless clearly indicates that the jump operator Q_0 needs extra treatment. For spline approximation of neutral functional differential equations this has turned out to be essential, both in theory and in numerical work [7].

For the rest of this subsection, we choose $\mu = 0$ in $(\mathscr{P}^{N,\mu})$ and let $N = 0,1,2,\ldots$.

We denote by $Z_1^N = \{(\phi(0),\phi) \in \mathscr{L} \mid \phi$ first order spline with knots at t_j^N, $j = 1,\ldots,N\}$, where $t_j^N = -j\frac{r}{N}$, $j = 0,\ldots,N$.

P_1^N stands for the orthogonal projection $Z_g \to Z_1^N$, $N = 1,\ldots$. It is proved in [3] that the approximation scheme $\{Z_1^N, P_1^N, P_1^N\mathscr{A}P_1^N\}$ satisfies the hypotheses of Theorem 3.1 and that $\dim(Z_1^N) = n(N+1)$. Therefore, for each z there exists a real-valued function $\bar{\rho}_1(N,z)$ such that

$$\lim_{N\to\infty} \bar{\rho}_1(N,z) = 0 \quad \text{and} \quad |T(t)z - T^N(t)z| \le \bar{\rho}_1(N,z).$$

This and the estimates to follow hold on the interval $[0,t^*]$. By Theorem 3.1 and the above inequality,(H1) is trivially satisfied. Using the triangle inequality, interpolating spline functions and [14, Theorem 2.4] it follows by a simple density argument that

$$(3.14) \qquad |P_1^N z - z| \le \tilde{\rho}_1(N,z), \text{ with } \lim_{\mu\to\infty} \tilde{\rho}_1(N,z) = 0,$$

so that (H2) is verified. Moreover for all z, the operators Q_1^N are chosen as

$$Q_1^N = P_1^N Q_0,$$

or in terms of their represenation

$$(Q_1^N)_j = P_1^N(e_j,0),$$

where e_j, $j = 1,\ldots,n$ stands for the n unit vectors in $\mathbb{R}^n$, and 0 for the zero function. (H3) (i) holds trivially and a short calculation gives

$$(3.16) \qquad ||Q_1^N - Q_0||_{\mathscr{L}(\mathbb{R}^n;z)} \le \sqrt{n} \max_{e_i} \tilde{\rho}_1(\mu,(e_i,0)),$$

and

$$||Q_1^N||_{\mathscr{L}(\mathbb{R}^n;z)} \leq 1.$$

Thus (H3) is verified.

Finally

$$(3.17) \qquad \tilde{\rho}(N,Q_1^N) = ||T(t)Q_1^N - T^N Q_1^N|| =$$

$$= ||T(t)P_1^N Q_0 - T(t)Q_0|| + ||T(t)Q_0 - T^N P_1^N Q_0|| \leq$$

$$\leq e^{\tilde{\omega}t^*}||P_1^N Q_0 - Q_0|| + \sqrt{n}\max_{e_i}(\bar{\rho}_1(N,(e_i,0)) +$$

$$+ e^{\tilde{\omega}t^*}\tilde{\rho}_1(N,(e_i,0))) \leq$$

$$\leq 2\, e^{\tilde{\omega}t^*}\sqrt{n}(\max_{e_i} \bar{\rho}_1(N,(e_i,0)) + \max_{e_i} \tilde{\rho}_1(N,(e_i,0))).$$

Estimates (3.14)-(3.17) are exactly those needed for the convergence results of control, state, payoff and Riccati operators in Section 2.

By (3.13) we know that on subspaces of Z, determined by $\mathrm{Dom}(\mathscr{A}^k)$, $k > 0$, $\bar{\rho}_1$ will actually go to zero with a rate given by convergence of the generators. But this is always at the expense of $\bar{\rho}_1$ not only depending on z but also on (at least) the L^2-norm of its second derivative. So even if we dispense of (H3)(i) for a moment, high order convergence of $||T(t)Q_0 - T^N(t)Q_0||$ to zero seems quite unlikely in the light of Remark 2.1.

(γ) Cubic spline functions

In this subsection the general results of Section 2 are used to discuss subspaces Z_3^N of Z given by

$$Z_3^N = \{(\phi(0),\phi) \in \mathscr{L}^2 \mid \phi \text{ is a cubic spline with knots at } t_j^N,\ j = 0,\dots N\},$$

where again $t_j^N = \frac{-r}{N} j$, $j = 0,\dots,N$, and $P_3^N\colon Z_g \to Z_3^N$ are the orthogonal projections. It is quite simple to verify that the approximation scheme $\{Z_3^N, P_3^N, \mathscr{A}_3^N\}$ with $\mathscr{A}_3^N = P_3^N \mathscr{A} P_3^N$ satisfies the conditions of Theorem 3.1 with $\dim Z_3^N = n(N+3)$ and that for $\mu = 0$ we can derive results similar to subsection (β). (H1) is therefore trivially satisfied. Here, however, we restrict our attention to the question of rate of convergence on subspaces of Z.

For $k = 1,2,\dots$ we introduce

$$\mathscr{D}^k = \{(\phi(0),\phi) \in \mathscr{W}^{k,2} \mid \phi \in W^{k+1,2},\ \phi^{(i)}(0) = L(\phi^{(i-1)}),\ i = 1,\dots,k\}.$$

Notice that $\mathscr{D}^k$ is the domain of the infinitesimal generator of the solution semigroup of the autonomous equation (2.3) if considered in the Banach space $\mathscr{W}^{k,2}$, (with its natural norm).

In particular, this implies that

$$\mathscr{D}^k \text{ is dense in } \mathscr{W}^{k,2},\ k = 1,2,\dots .$$

Moreover,

$$\text{if } z \in \mathscr{D}^k \text{ then } z \in \mathrm{Dom}(\mathscr{A}^{k+1}).$$

In [3] it is proved that for $\hat{\psi} = (\psi(0),\psi)$

$$(3.15) \qquad |T^N(t)\hat{\psi} - T(t)\hat{\psi}| \le \bar{\rho}(N,\hat{\psi}) = O(\frac{1}{N^3})$$

for $\hat{\psi} \in \mathscr{D}^5$, where $O(\frac{1}{N^3})$ depends on $\psi^{(4)}$, and from [14, Theorem 6.9] it follows that

$$(3.16) \qquad |P^N\hat{\psi}_1 - \hat{\psi}_1| \le O(\frac{1}{N^3}) \quad \text{for} \quad \psi_1 \in W^{3,2}.$$

Therefore, for $\hat{\psi} \in \mathscr{D}^6$

$$(3.17) \qquad \rho(N,\hat{\psi}) = O(\frac{1}{N^3}).$$

We define the operators approximating Q_0 by

$$\hat{Q}_3^N = ((Q_3^N(0),Q_3^N)_1,\ldots,(Q_3^N(0),Q_3^N)_n),$$

where Q_3^N is the $n \times n$ function-valued matrix

$$Q_3^N = \begin{bmatrix} s_3^N & & 0 \\ & \ddots & \\ 0 & & s_3^N \end{bmatrix};$$

s_3^N can be chosen very conveniently as the unique $C^2(-r,0;\mathbb{R})$ function, given by

$$\dot{s}^N(t_1^N) = \ddot{s}^N(t_1^N) = s^N(t_j^N) = 0 \quad \text{for} \quad j = 1,\ldots,N$$
$$s^N(t_0^N) = 1.$$

Notice that this choice of s_3 lets the diagonal of the representative of Q_3^N become a basis function of Z^N (possibly after multiplying with some scalar) in the commonly chosen de Boor basis.

The function s_3 can be explicitly represented as

$$s_3(t) = \begin{cases} (\frac{N}{r})^3 (t+\frac{r}{N})^3 & \text{for} \quad t \in [-\frac{r}{N}, 0] \\ 0 & \text{otherwise} , \end{cases}$$

and a short calculation yields

$$(3.18) \qquad \rho_Q(N) = ||Q_0 - Q_3^N||_{\mathscr{L}(\mathbb{R}^n, Z)} = O(\frac{1}{\sqrt{N}})$$

and

$$(3.19) \qquad ||Q_0 - P^N Q_0||_{\mathscr{L}(\mathbb{R}^n, Z)} = O(\frac{1}{\sqrt{N}}).$$

The special form of $\hat{Q}_3^N$ and (3.18) imply (H3)(i) and (ii). (H3)(iii) is verified easily. If $\mu + N$ is some multiple of N then (H4) holds and (H2) is a consequence of (3.16) and density of $\mathscr{S}^{3,2}$ in Z. Finally bounds on $\tilde{\rho}(N+\mu, Q^N)$ are given via the following lemmas. For $\hat{\phi} = (\phi(0), \phi) \in \mathscr{S}$ let ϕ_I^N denote the interpolating cubic spline function defined by

$$\phi_I^N(t_j^N) = \phi(t_j^N) \qquad \text{for} \quad j = 0, \dots, N$$

$$(\phi_I^N)'(0) = (\phi_I^N)'(-r) = 0.$$

Lemma 3.2. For all $\hat{\phi} \equiv (\phi(0),\phi) \in \mathscr{D}^2$ we have

$$(3.20) \qquad |\mathscr{A}_3^N\hat{\phi} - \mathscr{A}_3\hat{\phi}| = O(\tfrac{1}{N^2})|\phi^{(3)}|_{L^2},$$

where $O(\frac{1}{N^2})$ depends only on L and $\phi^{(3)}$ denotes the third derivate of ϕ.

Proof. Since P_3^N is an orthogonal projection

$$|P_3^N\mathscr{A}\hat{\phi} - \mathscr{A}\hat{\phi}|_{Z_g} = \min_{z\in Z_3^N}|z - \mathscr{A}\hat{\phi}|_{Z_g} \le \sqrt{\ell}\ |(\dot{\phi})_I^N - \dot{\phi}|_{L^2},$$

where we used the fact that $\phi \in \mathscr{D}^1$ and $\sqrt{\ell}$ is a consequence of the weighting function g. The last estimate, [14, Theorem 4.5] and $\phi \in \mathscr{D}^2$ imply

$$(3.21) \qquad |P_3^N\mathscr{A}\hat{\phi} - \mathscr{A}\hat{\phi}| = O(\tfrac{1}{N^2})|\phi^{(3)}|_{L^2}.$$

We now turn to estimate $\mathscr{A}P_3^N\hat{\phi} - \mathscr{A}\hat{\phi}$ and let $P_3^N\hat{\phi} = (\phi^N(0),\phi^N)$. Then

$$|(\phi^N-\phi)'|_{L^2} \le |(\phi^N-\phi_I^N)'|_{L^2} + |(\phi_I^N-\phi)'|_{L^2} \le$$

$$\le C_1N|\phi^N-\phi_I^N|_{L^2} + O(\tfrac{1}{N^2})|\phi^{(3)}|_{L^2},$$

where the first term is estimated by the Schmidt inequality [14, pg. 7], the second by [14, Theorem 6.9], and C_1 is a constant independent of N and ϕ.

The last inequality implies

$$(3.22)\qquad |D(\phi^N-\phi)|_{L^2} = O(\tfrac{1}{N^2})|\phi^{(3)}|_{L^2}.$$

A similar calculation gives

$$\sup_{s\in[-r,0]} |\phi(s) - \phi^N(s)| = O(\tfrac{1}{N^2})|\phi^{(3)}|_{L^2},$$

and therefore, since L is a bounded linear function from $C(-r,0;\mathbb{R}^n) \to \mathbb{R}^n$

$$(3.23)\qquad |L(\phi-\phi^N)| \le O(\tfrac{1}{N^2})|\phi^{(3)}|_{L^2}.$$

(3.21)-(3.23) are used in the final estimate

$$|\mathscr{A}_3^N\hat{\phi}-\mathscr{A}\hat{\phi}| \le |\mathscr{A}_3^N\hat{\phi}-P_3^N\mathscr{A}\hat{\phi}| + |P_3^N\mathscr{A}\hat{\phi}-\mathscr{A}\hat{\phi}| \le$$

$$\le |\mathscr{A}P_3^N\hat{\phi}-\mathscr{A}\hat{\phi}| + |P_3^N\mathscr{A}\hat{\phi}-\mathscr{A}\hat{\phi}| \le O(\tfrac{1}{N^2})|\phi^{(3)}|_{L^2},$$

which ends the proof.

<u>Lemma 3.3</u>. For all N, and $t \in [0,t^*]$

$$|(T^\mu(t) - T(t))Q_3^N|_{\mathscr{L}(\mathbb{R}^n;Z)} = O(\tfrac{1}{\mu^2})\ |s_3^N|_{W^{3,2}(-r,0;\mathbb{R})}$$

<u>Proof</u>. For an arbitrary j let $\hat{q}^N = (q^N(0),q^N) = (Q_3^N)_j$. Notice that $\hat{q}^N \in W^{3,2}(-r,0;\mathbb{R}^n)$, so that by density of $\mathscr{D}^3$ in

$W^{3,2}(-r,0;\mathbb{R}^n)$ there exists a sequence of functions $\hat{q}_n \in \mathscr{D}^3$ such that

$$\hat{q}_n \to \hat{q}^N \text{ in } W^{3,2}. \tag{3.24}$$

We turn to estimating $|T^\mu(t)\hat{q}_n - T(t)\hat{q}_n|$ first. Since $q_n \in W^{4,2}$ there exists a constant k_1, depending only on L and t^* such that

$$|T(t)\hat{q}_n|_{W^{3,2}} \leq k_1|q_n|_{W^{3,2}} \text{ for } t \in [0,t^*]. \tag{3.25}$$

Using the fact that $\hat{q}_n \in \mathscr{D}^3$, it is easy to check that $(\lambda_0 I-\mathscr{A})\hat{q}_n \in \mathscr{D}^2$, for some fixed $\lambda_0 > \tilde{\omega}$. Therefore, there exists another constant k_2, depending only on L and t^* such that

$$|T(t)(\lambda_0-\mathscr{A})\hat{q}_n|_{W^{3,2}} \leq k_2|q_n|_{W^{3,2}} \tag{3.26}$$

If we use (3.25) and (3.26) together with (3.20) in (3.13), we get

$$|T^\mu(t)\hat{q}_n - T(t)\hat{q}_n| = O(\tfrac{1}{\mu^2})|q_n|_{W^{3,2}},$$

where the $O(\frac{1}{\mu^2})$-term is independent of q_n and $t \in [0,t^*]$. The last estimate, together with

$$|T^\mu(t)\hat{q}^N - T(t)\hat{q}^N| \leq |T^\mu(t)\hat{q}^N - T^\mu(t)\hat{q}_n| + |T^\mu(t)\hat{q}_n - T(t)\hat{q}_n| +$$

$$+ |T(t)\hat{q}_n - T(t)\hat{q}^N| \leq 2e^{\tilde{\omega}t^*}|\hat{q}^N-\hat{q}_n| + O(\tfrac{1}{N^2})|\hat{q}^N|_{W^{3,2}} +$$

$$+ O(\tfrac{1}{N^2})|\hat{q}_n-\hat{q}^N|_{W^{3,2}},$$

which implies the claim.

The estimates (3.15)-(3.19) and Lemma 3.3 are exactly those estimates, which are needed to apply the results of Section 2, and essentially establish that cubic spline approximations to the linear-quadratic optimal control problem (P) are $O(\frac{1}{\mu^2})$ convergent for trajectories, controls and payoffs, if the initial data are chosen from certain subspaces.

4. The Averaging Approximation Scheme

When applying the results of Section 2 to averaging approximation schemes, the approximating state - and Riccati equations are found to be of particularly simple structure; moreover for the class of problems under consideration we find exactly those equations approximating problem (P) that were first proposed in [13] and [15]; they were also derived in [2], and convergence proofs for controls, state and payoff are provided in [1].

For any positive integer N we partition the interval $[-r,0]$ into the subintervals $[t_j^N, t_{j-1}^N]$ with $t_j^N = \frac{-jr}{N}$, for $j = 0,\dots,N$. Let χ_j^N denote the characteristic function of $[t_j^N, t_{j-1}^N)$ for $j = 2,\dots,N$ and χ_1^N is the characteristic function of $[t_1^N, t_0^N]$. Then the averaging approximation subspaces Z_{av}^N of Z are defined by

$$Z_{av}^N = \{(\eta,\phi) \mid \eta \in \mathbb{R}^n,\ \phi = \sum_{j=1}^N v_j^N \chi_j^N,\ v_j^N \in \mathbb{R}^n\}.$$

We note that $(\eta,0) \in Z_{av}^N$ for each $\eta \in \mathbb{R}^n$. It is simple to calculate the orthogonal projection $P_{av}^N : Z \to Z_{av}^N$; indeed for $(\eta,\phi) \in Z$ we have

$$P_{av}^N(\eta,\phi) = (\eta, \sum_{j=1}^N \phi_j^N \chi_j^N) \tag{4.1}$$

where $\phi_j^N = \frac{N}{r} \int_{t_j^N}^{t_{j-1}^N} \phi(s)ds$.

A scheme for approximating $T(t)$ using the subspaces Z_{av}^N has been derived in [1]. This "averaging approximation scheme" is described next. Again, we assume that the matrices A_i in (2.2) are independent of t and define a sequence of operators $\mathscr{A}^N_{av}: Z \to Z^N_{av}$ by

$$\mathscr{A}^N_{av}(\eta,\phi) = \Big(A_0\eta + \sum_{i=1}^{\ell}\sum_{j=1}^{N} A_i\phi_j^N\chi_j^N(-r_i) + \sum_{j=1}^{N}\frac{r}{N}D_j^N\phi_j^N,\ \sum_{j=1}^{N}\frac{N}{r}(\phi_{j-1}^N-\phi_j^N)\chi_j^N\Big)$$

where

$$\phi_0^N = \eta,\ \phi_j^N = \frac{N}{r}\int_{t_j^N}^{t_{j-1}^N}\phi(s)ds,\quad \text{and}\quad D_j^N = \frac{N}{r}\int_{t_j^N}^{t_{j-1}^N}A_{-1}(s)ds,\ j = 1,\dots,N.$$

It was shown in [1] that

$$(4.2)\quad \left\{\begin{array}{l} \text{if we fix the weight functions } \rho \equiv 1 \\ \mathscr{A}^N_{av} \text{ generates semigroups } T^N_{av}(t) \text{ such that} \\ \|T^N_{av}(t)\| \le M^* e^{\omega^* t} \text{ for } t \ge 0, \text{ with } M^* = M^*(A_i) \text{ and} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \omega^* = \omega^*(A_i) \\ T^N_{av} Z^N_{av} \subset Z^N_{av} \\ \lim_N P^N_{av} z = z, \text{ for all } z \in Z, \\ |T(t)z - T^N_{av}(t)z| \le \bar{\rho}_{av}(N,z) \text{ with } \lim_N \bar{\rho}_{av}(N,z) = 0, \\ \qquad\qquad \text{for } t \text{ in compact subsets of } [0,\infty) \end{array}\right.$$

so that (H1) and (H2) of Section 2 are satisfied. The operators Q^N are chosen as

$$(4.3)\qquad Q^N = P^N_{av}Q_0 = Q_0,$$

which, of course, implies that (H3) is trivially satisfied. By (14.1) we also have

$$(4.4) \qquad F^N = \mathscr{F}, \quad \mathscr{D}^N = \mathscr{D} \quad \text{and} \quad \mathscr{B}^N = \mathscr{B} = Q_0 B.$$

Now the estimates of Section 2 can be applied; for the optimal controls, for example, we get by Corollary 2.1

$$(4.5) \qquad |\tilde{u}-\tilde{u}^{N,0}|_{L^2(t_0,t^*;\mathbb{R}^m)} \le K_7^{av}[\tilde{\rho}_{av}(N,Q_0)|z| + \rho_{av}(N,z)],$$

and similarly

$$(4.6) \qquad |S(t)z - S_{av}^{N,0}(t)P_{av}^N z| \le K_8^{av}[\tilde{\rho}_{av}(N,Q_0)|z| + \rho_{av}(N,z)], \quad \text{for } t_0 \le t \le t^*,$$

$$(4.7) \qquad |\langle \Pi(t)z - \Pi^{N,0}(t)P^N z,y\rangle| \le K_9^{av}[\bar{\rho}_{av}(N,y)|z| + \rho_{av}(N,z)|y| + |y||z|\tilde{\rho}(N,Q_0)]$$

and by Theorem 2.2

$$(4.8) \qquad |J(t_0,P^N z,\tilde{u}^N) - J(t_0,z,\tilde{u})| \le K_{10}^{av}[\rho_{av}(N,z)|z| + \tilde{\rho}_{av}(N,Q_0)|z|^2],$$

so that in view of (4.2) convergence of optimal controls, optimal states and payoff as well as weak convergence of the Riccati operators is guaranteed.

Finally we give the form of the approximating state and Riccati equations. We use $e_0^N,\dots,e_N^N$ defined by

$$e_0^N = (1,0) \quad \text{and} \quad e_j^N = (0,\chi_j^N),\ j = 1,\ldots,N$$

as a basis for Z^N. Since $T_{av}^N(t)$ leaves Z_{av}^N invariant, (2.9) is equivalent to

$$(4.9) \qquad \begin{cases} \dot{z}^{N,0}(t) = A_{av}^N z^{N,0}(t) + (B(t)u^N(t),0) \\ z^{N,0}(t_0) = P_{av}^N(\eta,\phi), \end{cases}$$

which, in turn, is equivalent to

$$(4.10) \qquad \begin{cases} \dot{w}^N(t) = A_{av}^N w^N(t) + \mathrm{col}(Bu^N(t),0,\ldots,0) \\ w^N(t_0) = \mathrm{col}(\eta,\phi_1^N,\ldots,\phi_N^N), \end{cases}$$

where $z^{N,0}(t) = \sum_{j=0}^{N} w_j^N(t)e_j^N$, ϕ_j^N is defined in (4.1), and

$$A_{av}^N = \begin{bmatrix} A_0 & \frac{r}{N}D_1^N & 0 & \cdots & 0 & \frac{r}{N}D_{N-1}^N & A_\ell + \frac{r}{N}D_N^N \\ \frac{N}{r}I & -\frac{N}{r}I & 0 & \cdots & & 0 & 0 \\ 0 & \frac{N}{r}I & -\frac{N}{r}I & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & & & \vdots \\ & \vdots & & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & & 0 & \frac{N}{r}I & -\frac{N}{r}I \end{bmatrix};$$

here I is the $n \times n$ identity matrix. If we let π_{av}^N denote the matrix representation of $\Pi^{N,0}: Z^N \to Z^N$, then we find by (2.17) that π_{av}^N satisfies an $[t_0,t^*]$ the matrix Riccati equation

$$\frac{d}{dt}\pi^N_{av}(t) = -(A^N_{av})^*\pi^N_{av}(t) - \pi^N_{aw}(t)A^N_{av} - [D]$$

(4.11)

$$+ \pi^{N(t)}_{av}[B(t)]C^{-1}[B(t)]^*\pi^N_{av}(t)$$

$$\pi^N_{av}(t) = [F],$$

where $(A^N_{av})^*$ denotes the transpose of A^N_{av} and

$$[D] = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \quad [F] = \begin{pmatrix} F & 0 \\ 0 & 0 \end{pmatrix}, \quad [B(t)] = \mathrm{col}(B(t),0,\ldots,0),$$

$[D]$ and $[F]$ being of dimension $n(N+1) \times n(N+1)$ and $[B(t)]$ of dimension $n(N+1) \times n$.

The optimal feedback law becomes

$$u^N(t) = -C^{-1}[B(t)]^*\pi^N_{av}(t)w^N(t). \tag{4.12}$$

Equations (4.10)-(4.12) completely describe the approximating linear regulator problem on the finite interval $[t_0,t^*]$.

5. Proofs.

In this section we give the proofs of those results that were not verified in Section 2. In addition to the conventions specified there, we let $P^N z = z^N$ for $z \in Z$ and $N = 1,2,\ldots$.

Proof of Lemma 4.

(a) We use (2.12) to find the following estimate for $W_{t_0}^{N,\mu}$:

$$|W^{N,\mu}z(t)|_{\mathbb{R}^m} = |(\mathscr{B}^N(t))^* \int_t^{t^*} (T^{N+\mu}(\eta,t))^* \mathscr{D}^{N+\mu}(\eta) T^{N+\mu}(\eta,t_0) z \, d\eta| +$$

$$+ |(\mathscr{B}^N(t))^* (T^{N+\mu}(t^*,t))^* F^{N+\mu} T^{N+\mu}(t^*,t_0) z| \leq$$

$$\leq bq(t^*-t_0)\overline{M}^2 e^{2\overline{\omega}(t^*-t_0)} d|z| +$$

$$+ bqf\overline{M}^2 e^{2\overline{\omega}(t^*-t_0)}|z| = bq\overline{M}^2 e^{2\overline{\omega}(t^*-t_0)}|z|(d(t^*-t_0) + f).$$

(b) The conditions on C,D,F and (2.11) imply (b) after a short calculation.

(c) For $t \in [t_0,t^*]$ we have

$$|(W_0 z)(t) - (W_0^{N,\mu} z^{N+\mu})(t)|_{\mathbb{R}^m} = |\mathscr{B}(t)^* \int_t^{t^*} T^*(\eta,t)\mathscr{D}(\eta)T(\eta,t_0)z \, d\eta -$$

$$- \mathscr{B}^N(t)^* \int_t^{t^*} T^{N+\mu}(\eta,t)^* \mathscr{D}^{N+\mu}(\eta) T^{N+\mu}(\eta,t_0) z^{N+\mu} d\eta| +$$

$$+ |\mathscr{B}(t)^* T(t^*,t)^* \mathscr{F} T(t^*,t_0) z -$$

$$- \mathscr{B}^N(t)^* T^{N+\mu}(t^*,t)^* F^{N+\mu} T^{N+\mu}(t^*,t_0) z^{N+\mu}| =$$

$$= |I_W(t)| + |II_W(t)|.$$

The terms I_w and II_w are now estimated separately. Let $\tau(t) = \operatorname{sgn} I_w(t)$, then

$$|I_w(t)|_{\mathbb{R}^m} \leq (I_w(t),\tau(t))_{\mathbb{R}^m} = \left(\int_t^{t^*} T(\eta,t)^* \mathscr{D}(\eta) T(\eta,t_0) z\, dy,\ \mathscr{B}(t)\tau(t)\right) -$$

$$- \left(\int_t^{t^*} T^{N+\mu}(\eta,t)^* \mathscr{D}^{N+\mu}(\eta) T^{N+\mu}(\eta,t_0) z^{N+\mu} d\eta, \mathscr{B}^N(t)\tau(t)\right) =$$

$$= \left(\int_t^{t^*} T(\eta,t)^* \mathscr{D}(\eta) T(\eta,t_0) z\, d\eta,\ (\mathscr{B}(t) - \mathscr{B}^N(t))\tau(t)\right) +$$

$$+ \left(\int_t^{t^*} T(\eta,t)^* \mathscr{D}(\eta) T(\eta,t_0) z - \int_t^{t^*} T^{N+\mu}(\eta,t)^* \mathscr{D}(\eta) T^{N+\mu}(\eta,t_0) z^{N+\mu} d\eta,\ \mathscr{B}^N(t)\tau(t)\right) +$$

$$+ \left(\int_t^{t^*} T^{N+\mu}(\eta,t)^* (\mathscr{D}(\eta) - \mathscr{D}^{N+\mu}(\eta)) T^{N+\mu}(\eta,t_0) z^{N+\mu} d\eta, \mathscr{B}^N(t)\tau(t)\right) \leq$$

$$\leq d\bar{M}^2 \int_t^{t^*} e^{\bar{\omega}(\eta-t)} e^{\bar{\omega}(\eta-t_0)} |z|\, d\eta\ b\ \rho_Q(N)\sqrt{n} +$$

$$+ \int_t^{t^*} \langle \mathscr{D}(\eta)(T(\eta,t_0)z - T^{N+\mu}(\eta,t_0)z^{N+\mu}),\ T(\eta,t)\mathscr{B}^N(t)\tau(t)\rangle_Z d\eta +$$

$$+ \int_t^{t^*} \langle \mathscr{D}(\eta) T^{N+\mu}(\eta,t_0) z^{N+\mu},\ T(\eta,t)\mathscr{B}^N(t)\tau(t) - T^{N+\mu}(\eta,t)\mathscr{B}^N(t)\tau(t)\rangle_Z d\eta +$$

$$+ \int_t^{t^*} \langle (\mathscr{D}(\eta) - \mathscr{D}^{N+\mu}(\eta)) T^{N+\mu}(\eta,t_0) z^{N+\mu},\ T^{N+\mu}(\eta,t)\mathscr{B}^N(t)\tau(t)\rangle_Z d\eta \leq$$

$$\leq b\ d\sqrt{n}\ \bar{M}^2 |z|\ \rho_Q(N)(e^{2\bar{\omega}(t^*-t_0)} - 1)(2\bar{\omega})^{-1} +$$

$$+ bqd\sqrt{n}\ \bar{M}(e^{\bar{\omega}(t^*-t_0)} - 1)\bar{\omega}^{-1}\,[\rho(N+\mu,z) + \tilde{\rho}(N+\mu,Q^N)|z|] +$$

$$+ bqd\sqrt{n}\ \bar{M}^2 (t^*-t_0) e^{2\bar{\omega}(t^*-t_0)} (2\bar{\omega})^{-1} |z|\ ||P^{N+\mu}Q_0 - Q_0||.$$

Let $e(t) = \operatorname{sgn} II_w(t)$, then

$$|II_w(t)| \leq (II_w(t),e(t))_{\mathbb{R}^m} \leq$$

$$\leq (\mathscr{B}(t)^* T(t^*,t)^* \mathscr{F} T(t^*,t_0)z - \mathscr{B}^N(t)^* T(t^*,t)^* \mathscr{F} T(t^*,t_0)z,\ e(t)) +$$

$$+ (\mathscr{B}^N(t)^* T(t^*,t)^* \mathscr{F} T(t^*,t_0)z - \mathscr{B}^N(t)^* T^{N+\mu}(t^*,t)^* \mathscr{F} T^{N+\mu}(t^*,t_0)z^{N+\mu},e(t)) +$$

$$+ (\mathscr{B}^N(t)^* T^{N+\mu}(t^*,t)^* (\mathscr{F} T^{N+\mu}(t^*,t_0)z^{N+\mu} - F^{N+\mu} T^{N+\mu}(t^*,t_0)z^{N+\mu},\ e(t)) =$$

$$= \langle T^*(t^*,t)\mathscr{F} T(t^*,t)z,\ (\mathscr{B}(t) - \mathscr{B}^N(t))e(t)\rangle +$$

$$+ \langle \mathscr{F} T(t^*,t_0)z,\ T(t^*,t)\mathscr{B}^N(t)e(t)\rangle - \langle \mathscr{F} T^{N+\mu}(t^*,t_0)z^{N+\mu},\ T^{N+\mu}(t^*,t)\mathscr{B}^N(t)e(t)\rangle +$$

$$+ \langle (\mathscr{F} - P^{N+\mu}\mathscr{F})T^{N+\mu}(t^*,t_0)z^{N+\mu},\ T^{N+\mu}(t^*,t_0)\mathscr{B}^N(t)e(t)\rangle =$$

$$= bf\sqrt{n}\ |z|\ \bar{M} e^{2\bar{\omega}(t^*-t_0)} \rho_Q(N) +$$

$$+ \langle \mathscr{F} T(t^*,t_0)z - \mathscr{F} T^{N+\mu}(t^*,t_0)z^{N+\mu},\ T(t^*,t)\mathscr{B}^N(t)e(t)\rangle +$$

$$+ \langle \mathscr{F} T^{N+\mu}(t^*,t_0)z^{N+\mu},\ (T(t^*,t) - T^{N+\mu}(t^*,t))\mathscr{B}^N(t)e(t)\rangle +$$

$$+ \| \mathscr{F} - P^{N+\mu}\mathscr{F} \|\ \bar{M}^2 e^{2\bar{\omega}(t^*-t_0)} bq\sqrt{n}\ |z| \leq$$

$$\leq bf\sqrt{n}|z|\ \bar{M}^2 e^{2\bar{\omega}(t^*-t_0)} \rho_Q(N) +$$

$$+ bq\sqrt{n}|f|\bar{M}\ e^{\bar{\omega}(t^*-t_0)} [\rho(N+\mu,z) + |z|\tilde{\rho}(N+\mu,Q^N)] +$$

$$+ bq\sqrt{n}|z|\bar{M}^2 e^{2\bar{\omega}(t^*-t_0)} f\|P^{N+\mu}Q_0 - Q_0\|.$$

The two inequalities together imply

$$\sup_{t\in[t_0,t^*]} (|I_w(t)| + |II_w(t)|) \leq$$

$$\rho_Q(N) b\sqrt{n}\ \bar{M}^2|z|e^{2\bar{\omega}(t^*-t_0)} [(2\bar{\omega})^{-1}d + f] + \tilde{\rho}(N+\mu,Q^N) bq\sqrt{n}\ \bar{M}|z|[d(e^{\bar{\omega}(t^*-t_0)}-1)\bar{\omega}^{-1} +$$

$$+ f\ e^{\bar{\omega}(t^*-t_0)}] + \rho(N+\mu,z) bq\sqrt{n}\ \bar{M}[d(e^{\bar{\omega}(t^*-t_0)}-1)\bar{\omega}^{-1} + f\ e^{\bar{\omega}(t^*-t_0)}] +$$

$$+ \|P^{N+\mu}Q_0 - Q_0\| bq\sqrt{n}\ \bar{M}^2|z|[d(t^*-t_0)e^{2\bar{\omega}(t^*-t_0)}(2\bar{\omega})^{-1} + f\ e^{2\bar{\omega}(t^*-t_0)}].$$

(d) For $w \in L^2(t_0,t^*;\mathbb{R}^m)$ we next estimate

$$|((V_0-V_0^{N,\mu})w)(t)| =$$

$$= |\mathscr{B}(t)^*T(t^*,t)^*\mathscr{F}\int_t^{t^*}T(t^*,\sigma)\mathscr{B}(\sigma)w(\sigma)d\sigma -$$

$$- \mathscr{B}^N(t)^*T^{N+\mu}(t^*,t)^*F^{N+\mu}\int_{t_0}^{t^*}T^{N+\mu}(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma| +$$

$$+ |\mathscr{B}(t)^*\int_t^{t^*}T(\eta,t)^*\mathscr{D}(\eta)\int_{t_0}^{t}T(\eta,\sigma)\mathscr{B}(\sigma)w(\sigma)d\sigma d\eta -$$

$$- \mathscr{B}^N(t)^*\int_t^{t^*}T^{N+\mu}(\eta,t)^*\mathscr{D}^{N+\mu}(\eta)\int_{t_0}^{\eta}T^{N+\mu}(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma| =$$

$$= |I_v(t)| + |II_v(t)|.$$

Again, we estimate I_v and II_v separately. We let $e(t) = \operatorname{sgn} I_v(t)$, then

$$|I_v(t)| = (\mathscr{B}(t)^*T(t^*,t)^*\mathscr{F}\int_{t_0}^{t^*}T(t^*,\sigma)\mathscr{B}(\sigma)w(\sigma)d\sigma -$$

$$- \mathscr{B}^N(t)^*T(t^*,t)^*\mathscr{F}\int_{t_0}^{t^*}T(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ e(t)) +$$

$$+ (\mathscr{B}^N(t)^*T(t^*,t)^*\mathscr{F}\int_{t_0}^{t^*}T(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ e(t)) -$$

$$- (\mathscr{B}^N(t)^*T^{N+\mu}(t^*,t)^*\mathscr{F}\int_{t_0}^{t^*}T^{N+\mu}(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ e(t)) +$$

$$+ (\mathscr{B}^N(t)^*T^{N+\mu}(t^*,t)^*(\mathscr{F}-F^{N+\mu})\int_{t_0}^{t^*}T^{N+\mu}(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ e(t)) =$$

$$= \langle T(t^*,t)^*\mathscr{F}\int_{t_0}^{t^*}T(t^*,\sigma)\mathscr{B}(\sigma)w(\sigma)d\sigma,\ (\mathscr{B}(t) - \mathscr{B}^N(t))e(t)\rangle +$$

$$+ \langle T(t^*,t)^* \mathscr{F} \int_{t_0}^{t^*} T(t^*,\sigma)(\mathscr{B}(\sigma) - \mathscr{B}^N(\sigma))w(\sigma)d\sigma,\ \mathscr{B}^N(t)e(t)\rangle +$$

$$+ \langle \mathscr{F} \int_{t_0}^{t^*} T(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ T(t^*,t)\mathscr{B}^N(t)e(t)\rangle -$$

$$- \langle \mathscr{F} \int_{t_0}^{t^*} T^{N+\mu}(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ T^{N+\mu}(t^*,t)\mathscr{B}^N(t)e(t)\rangle +$$

$$+ \langle(\mathscr{F}-F^{N+\mu}) \int_{t_0}^{t^*} T^{N+\mu}(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ T^{N+\mu}(t^*,t)\mathscr{B}^N(t)e(t)\rangle \le$$

$$\le b^2 f\sqrt{n}\,\bar{M}^2\ (1+q)\rho_Q(N)e^{\bar{\omega}(t^*-t)}[\frac{1}{2\bar{\omega}}\ (e^{2\bar{\omega}(t^*-t_0)}-1)]^{1/2}|w|_{L^2} +$$

$$+ \langle \mathscr{F} \int_{t_0}^{t^*} T(t^*,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ (T(t^*,t) - T^{N+\mu}(t^*,t))\mathscr{B}^N(t)e(t)\rangle +$$

$$+ \langle \mathscr{F} \int_{t_0}^{t^*} (T(t^*,\sigma) - T^{N+\mu}(t^*,\sigma))\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ T^{N+\mu}(t^*,t)\mathscr{B}^N(t)e(t)\rangle +$$

$$+ ||\mathscr{F}-P^{N+\mu}\mathscr{F}||b^2q^2\sqrt{n}\,\bar{M}^2 e^{\bar{\omega}(t^*-t_0)}\left[\frac{e^{2\bar{\omega}(t^*-t_0)}-1}{2\bar{\omega}}\right]^{1/2} |w|_{L^2} \le$$

$$\le b^2 f\sqrt{n}\,\bar{M}^2(1+q)\rho_Q(N)e^{\bar{\omega}(t^*-t)}\left[\frac{1}{2\bar{\omega}}\ (e^{2\bar{\omega}(t^*-t_0)}-1\right]^{1/2} |w|_{L^2} +$$

$$+ \bar{M}b^2qf\sqrt{n}\ \tilde{\rho}(N+\mu,Q^N)|w|_2\left[\left(\frac{e^{2\bar{\omega}(t^*-t_0)}-1}{2\bar{\omega}}\right)^{1/2} + e^{\bar{\omega}(t^*-t_0)}(t^*-t_0)^{1/2}\right] +$$

$$+ ||Q_0 - P^{N+\mu}Q_0||fb^2q^2\sqrt{n}\,\bar{M}^2 e^{\bar{\omega}(t^*-t_0)}\left[\frac{e^{2\bar{\omega}(t^*-t_0)}-1}{2\bar{\omega}}\right]^{1/2} |w|_{L^2}.$$

Finally, with $v(t) = \operatorname{sgn} II_v(t)$ we get

$$|II_v(t)| = (II_v(t), v(t)) =$$

$$= \Big(\mathscr{B}(t)^* \int_t^{t^*} T(\eta,t)^* \mathscr{D}(\eta) \int_{t_0}^{\eta} T(\eta,\sigma)\mathscr{B}(\sigma)w(\sigma)d\sigma d\eta -$$

$$- \mathscr{B}^N(t)^* \int_t^{t^*} T^{N+\mu}(\eta,t)^* \mathscr{D}^{N+\mu}(\eta) \int_{t_0}^{\eta} T^{N+\mu}(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma d\eta,\ v(t)\Big) =$$

$$= \Big(\mathscr{B}(t)^* \int_t^{t^*} T(\eta,t)^* \mathscr{D}(\eta) \int_{t_0}^{\eta} T(\eta,\sigma)\mathscr{B}(\sigma)w(\sigma)d\sigma d\eta -$$

$$- \mathscr{B}^N(t)^* \int_t^{t^*} T(\eta,t)^* \mathscr{D}(\eta) \int_{t_0}^{\eta} T(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma d\eta,\ v(t)\Big) +$$

$$+ \Big(\mathscr{B}^N(t)^* \int_t^{t^*} T(\eta,t)^* \mathscr{D}(\eta) \int_{t_0}^{\eta} T(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma d\eta -$$

$$- \mathscr{B}^N(t)^* \int_t^{t^*} T^{N+\mu}(\eta,t)^* \mathscr{D}(\eta) \int_{t_0}^{\eta} T^{N+\mu}(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)\, d\sigma d\eta, v(t)\Big) +$$

$$+ \Big(\mathscr{B}^N(t)^* \int_t^{t^*} T^{N+\mu}(\eta,t)^* (\mathscr{D}(\eta) -$$

$$- \mathscr{D}^{N+\mu}(\eta)) \int_{t_0}^{\eta} T^{N+\mu}(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma d\eta,\ v(t)\Big) =$$

$$= <\int_t^{t^*} T^*(\eta,t)\mathscr{D}(\eta) \int_{t_0}^{\eta} T(\eta,\sigma)(\mathscr{B}(\sigma) - \mathscr{B}^N(\sigma))w(\sigma)d\sigma d\eta,\ \mathscr{B}(t)v(t)> +$$

$$+ <\int_t^{t^*} T^*(\eta,t)\mathscr{D}(\eta) \int_{t_0}^{\eta} T(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma d\eta,\ (\mathscr{B}(t) - \mathscr{B}^N(t)v(t)> +$$

$$+ \int_t^{t^*} (<\mathscr{D}(\eta) \int_{t_0}^{\eta} (T(\eta,\sigma) - T^{N+\mu}(\eta,\sigma))\mathscr{B}^N(\sigma)w(\sigma)d\sigma,\ T^{N+\mu}(\eta,t)\mathscr{B}^N(t)v(t)>)d\eta +$$

$$+ <\mathscr{D}(\eta) \int_{t_0}^{\eta} (T(\eta,\sigma) - T^{N+\mu}(\eta,\sigma))\mathscr{B}^N(\sigma)w(\sigma)d\sigma,$$

$$T^{N+\mu}(\eta,t)\mathscr{B}^N(t)\,v(t)>)d\eta +$$

$$+ \int_t^{t^*} <(\mathscr{D}(\eta) - \mathscr{D}^{N+\mu}(\eta)) \int_{t_0}^{\eta} T^{N+\mu}(\eta,\sigma)\mathscr{B}^N(\sigma)w(\sigma)d\sigma,$$

$$T^{N+\mu}(\eta,t)\mathscr{B}^N(t)v(t)>d\eta \leq$$

$$\leq b^2 d\sqrt{n}\,\bar{M}^2\rho_Q(N)|w|_{L^2} \frac{1}{\sqrt{2\bar{\omega}^3}} [e^{2\bar{\omega}(t^*-t_0)}-1](1+q) +$$

$$+ \int_t^t d\left(\int_{t_0}^{\eta} \bar{M}\, e^{\bar{\omega}(\eta-\sigma)} bq|w(\sigma)|d\sigma\right)\tilde{\rho}(N+\mu,Q^N)bq\sqrt{n}\, d\eta +$$

$$+ \int_t^{t^*} d\left(\int_{t_0}^{\eta} \tilde{\rho}(N+\mu,Q^N)bq|w(\sigma)|d\sigma\right)\bar{M}e^{\bar{\omega}(\eta-t)}bq\sqrt{n}\, d\eta +$$

$$+ d||Q_0-P^{N+\mu}Q_0|| \int_t^{t^*}\int_{t_0}^{\eta} \bar{M}^2 e^{\bar{\omega}(\eta-\sigma)}b^2q^2|w(\sigma)|d\sigma\, e^{\bar{\omega}(\eta-t)}\sqrt{n}\, d\eta \leq$$

$$\leq b^2 d\sqrt{n}\,\bar{M}^2\, \rho_Q(N)|w|_{L^2} \frac{1}{\sqrt{2\bar{\omega}^3}} [e^{2\bar{\omega}(t^*-t_0)}-1](1+q) +$$

$$+ db^2q^2\bar{M}\sqrt{n}\, \tilde{\rho}(N+\mu,Q^N)|\int_t^{t^*}\int_{t_0}^{\eta} e^{\bar{\omega}(\eta-\sigma)}|w(\sigma)|d\sigma d\eta + \int_t^{t^*}\int_{t_0}^{\eta} e^{\bar{\omega}(\eta-t)}|w(\sigma)|d\sigma d\eta$$

$$+ db^2g^2\bar{M}^2\sqrt{n}||Q_0-P^{N+\mu}Q_0|| \int_t^{t^*}\int_{t_0}^{\eta} e^{\bar{\omega}(\eta-\sigma)}e^{\bar{\omega}(\eta-t)}|w(\sigma)|d\sigma d\eta \leq$$

$$\leq b^2 d\sqrt{n}\,\bar{M}^2\rho_Q(N)|w|_{L^2} \frac{1}{\sqrt{2\bar{\omega}^3}} [e^{2\bar{\omega}(t^*-t_0)}-1](1+q) +$$

$$+ \; db^2q^2\overline{M}\sqrt{n}\tilde{\rho}(N+\mu,Q^N)|w|_{L^2}\left[\frac{1}{\sqrt{2\overline{\omega}^3}}\,e^{\overline{\omega}(t^*-t_0)} + \frac{(t^*-t_0)^{1/2}}{\overline{\omega}}\,(e^{\overline{\omega}(t^*-t_0)}-1)\right]$$

$$+ \; db^2q^2\overline{M}^2\sqrt{n}\;||Q_0-P^{N+\mu}Q_0||\;\;|w|_{L^2}\left[\frac{1}{\sqrt{2\overline{\omega}^3}}\,(e^{2\overline{\omega}(t^*-t_0)}-1)\right].$$

The bounds on I_V and II_V imply

$$\sup_{t\in[t_0,t^*]}|((V_0-V_0^{N,\mu})w)(t)| \le$$

$$\le \rho_Q(N)b^2\sqrt{n}\,\overline{M}^2(1+q)|w|_{L^2}\left(\frac{f}{\sqrt{2\overline{\omega}}} + \frac{d}{\sqrt{2\overline{\omega}^3}}\right)e^{2\overline{\omega}(t^*-t_0)} +$$

$$+ \; \tilde{\rho}(N+\mu,Q^N)b^2\sqrt{n}\,\overline{M}\,q^2|w|_{L^2}\left(f(\frac{1}{\sqrt{2\overline{\omega}}} + (t^*-t_0)^{1/2}) + \right.$$

$$\left. + \; d(\frac{1}{\sqrt{2\overline{\omega}^3}} + \frac{(t^*-t_0)^{1/2}}{\overline{\omega}})\right)e^{\omega(t^*-t_0)}$$

$$+ \; ||Q_0-P^{N+\mu}Q_0||b^2q^2\overline{M}^2\sqrt{n}\;|w|_{L^2}\left[\frac{d}{\sqrt{2\overline{\omega}^3}}\,(e^{2\overline{\omega}(t^*-t_0)}-1) + \frac{f}{\sqrt{2\overline{\omega}}}\,e^{\overline{\omega}(t^*-t_0)}\right].$$

This completes the proof.

<u>Proof of Theorem 2.2.</u>

(a) $$|S^{N,\mu}(t,s)z^{N+\mu}| = |T^{N+\mu}(t,s)z^{N+\mu}| + |\int_s^t T^{N+\mu}(t,\eta)\mathscr{B}^N(\eta)((V_s^{N+\mu})^{-1}W_s^{N,\mu}z^{N+\mu})(\eta)d\eta|$$

$$\le \overline{M}e^{\overline{\omega}(t^*-t_0)}|z^{N+\mu}| + \overline{M}e^{\overline{\omega}(t^*-t_0)}bq\int_s^t|((V_s^{N,\mu})^{-1}W_s^{N,\mu}z^{N+\mu})(\eta)|d\eta \le$$

$$\leq \bar{M}e^{\bar{\omega}(t^*-t_0)}|z^{N+\mu}| + \bar{M}e^{\bar{\omega}(t^*-t_0)}bq\,(t^*-t_0)^{1/2}\,c^{-1}\,(t^*-t_0)^{1/2}\,K_1|z^{N+\mu}| =$$

$$= \bar{M}e^{\bar{\omega}(t^*-t_0)}|z|(1+bq(t^*-t_0)c^{-1}K_1).$$

(b) $$|S(t,s)z - S^{N,\mu}(t,s)P^{N+\mu}z| \leq |T(t,s)z - T^{N+\mu}(t,s)z^{N+\mu}| +$$

$$+ \int_s^t |T(t,\eta)\mathcal{B}(\eta)(V_s^{-1}W_s z)(\eta) - T(t,\eta)\mathcal{B}(\eta)((V_s^{N,\mu})^{-1}W_s^{N,\mu}z^{N+\mu})(\eta)|d\eta \leq$$

$$+ \int_s^t |T(t,\eta)\mathcal{B}(\eta)((V_s^{N,\mu})^{-1}W_s^{N,\mu}z^{N+\mu})(\eta) - T^{N+\mu}$$

$$- T^{N+\mu}(t,\eta)\mathcal{B}^N(\eta)((V_s^{N,\mu})^{-1}W_s^{N,\mu}z^{N+\mu})(\eta))d\eta \leq$$

$$\leq \rho(N+\mu,z) + \int_s^t \bar{M}e^{\bar{\omega}(t-\eta)}b|(V_s^{-1}W_s z)(\eta) - ((V_s^{N,\mu})^{-1}W_s^{N,\mu}z^{N+\mu})(\eta)|d\eta +$$

$$+ b\left(\int_s^t ||T(t,\eta)Q_0 - T^{N+\mu}(t,\eta)Q^N||^2 d\eta\right)^{1/2} \left(\int_s^t |((V_s^{N,\mu})^{-1}W_s^{N,\mu}z^{N+\mu})(\eta)|^2 d\eta\right)^{1/2} \leq$$

$$\leq \rho(N+\mu,z) + b\,|\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}\,\bar{M}\left(\int_s^t e^{2\bar{\omega}(t-\eta)}d\eta\right)^{1/2} +$$

$$+ b\left[\left(\int_s^t ||T(t,\eta)Q_0 - T(t,\eta)Q^N||^2\right)\eta^{\;1/2} +\right.$$

$$\left. + \left(\int_s^t ||T(t,\eta)Q^N - T^{N+\mu}(t,\eta)Q^N||^2 d\eta\right)^{1/2}\right] c^{-1}|z|\;K_1\;(t^*-t_0)^{1/2} \leq$$

$$\leq \rho(N+\mu,z) + b|\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}\bar{M}\,\left(\frac{1}{2\bar{\omega}}\,(e^{2\bar{\omega}(t^*-t_0)}-1)\right)^{1/2} +$$

$$+ bc^{-1}|z|K_1(t^*-t_0)^{1/2}\left[\bar{M}\rho_Q(N)\left(\int_s^t e^{2\bar{\omega}(t-\eta)}d\eta\right)^{1/2} + \tilde{\rho}(N+\mu,Q^N)(t^*-t_0)^{1/2}\right]$$

$$= \rho(N+\mu,z) + b|\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}\bar{M}\left(\frac{1}{2\bar{\omega}}\left(e^{2\bar{\omega}(t^*-t_0)}-1\right)\right)^{1/2} +$$

$$+ bc^{-1}|z|K_1(t^*-t_0)^{1/2}\bar{M}\rho_Q(N)\left(\frac{1}{2\bar{\omega}}\left(e^{2\bar{\omega}(t^*-t_0)}-1\right)\right)^{1/2} +$$

$$+ bc^{-1}|z|K_1(t^*-t_0)\ \tilde{\rho}(N+\mu,Q^N) =$$

$$= \rho(N+\mu,z) + \tilde{\rho}(N+\mu,Q^N)|z|K_6 + |\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}k_5 + |z|\rho_Q(N)k_6,$$

where

$$K_6 = bc^{-1}K_1(t^*-t_0)$$

$$k_5 = b\bar{M}\left(\frac{1}{2\bar{\omega}}\left(e^{2\bar{\omega}(t^*-t_0)}-1\right)\right)^{1/2}$$

$$k_6 = bc^{-1}K_1(t^*-t_0)^{1/2}\bar{M}\left(\frac{1}{2\bar{\omega}}\left(e^{2\bar{\omega}(t^*-t_0)}-1\right)\right)^{1/2}.$$

<u>Proof of Theorem 2.3.</u>

(a) By (2.13) we have for $t \in [t_0,t^*]$

$$|\Pi^{N,\mu}(t)z^{N+\mu}| \le |T^{N+\mu}(t^*,t)^*F^{N+\mu}S^{N,\mu}(t^*,t)z^{N+\mu}| +$$

$$+ \left|\int_t^{t^*} T^{N+\mu}(\eta,t)^*\mathcal{D}^{N+\mu}(\eta)S^{N,\mu}(\eta,t)z^{N+\mu}d\eta\right| \le$$

$$\le \bar{M}e^{\bar{\omega}(t^*-t_0)}fK_5|z| + \bar{M}e^{\bar{\omega}(t^*-t_0)}dK_5|z|(t^*-t_0).$$

(b) $$<\Pi(t)z - \Pi^{N,\mu}(t)z^{N+\mu},y> = <T(t^*,t)^*\mathcal{F}S(t^*,t)z -$$

$$- T^{N+\mu}(t^*,t)^*F^{N+\mu}S^{N,\mu}(t^*,t)z^{N+\mu},y> +$$

$$+ \int_t^{t^*} <T(\eta,t)^*\mathcal{D}(\eta)S(\eta,t)z -$$

$$- T^{N+\mu}(\eta,t)^*\mathcal{D}^{N+\mu}(\eta)S^{N,\mu}(\eta,t)z^{N+\mu},y>d\eta \le$$

$$\le <\mathcal{F}S(t^*,t)z,\ T(t^*,t)y - T^{N+\mu}(t^*,t)y> + <\mathcal{F}S(t^*,t)z -$$

$$- <\mathcal{F}S^{N,\mu}(t^*,t)z^{N+\mu},T^{N+\mu}(t^*,t)y> +$$

$$+ <(\mathcal{F}-F^{N+\mu})S^{N,\mu}(t^*,t)z^{N+\mu},T^{N+\mu}(t^*,t)y> +$$

$$+ \int_t^{t^*} (<\mathcal{D}(\eta)S^{N,\mu}(\eta,t)z,\ T(\eta,t)y - T^{N+\mu}(\eta,t)y> +$$

$$+ <\mathcal{D}(\eta)S(\eta,t)z - \mathcal{D}(\eta)S^{N,\mu}(\eta,t)z^{N+\mu},T^{N+\mu}(\eta,t)y>)d\eta$$

$$+ \int_t^{t^*} <\mathcal{D}(\eta)S^{N,\mu}(\eta,t)z^{N+\mu} - \mathcal{D}^{N+\mu}(\eta)S^{N,\mu}(\eta,t)z^{N+\mu},T^{N+\mu}(\eta,t)y>d\eta \le$$

$$\le fK_5|z|\bar{\rho}(N+\mu,y) + (f+d(t^*-t_0))[Me^{\bar{\omega}(t^*-t_0)}|y|\,(\rho(N+\mu,z) +$$

$$+ \tilde{\rho}(N+\mu,Q^N)|z|K_6 + |\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}k_5 + |z|\rho_Q(N)k_6] +$$

$$+ \|\,Q_0-P^{N+\mu}Q_0\,\|fK_5|z|Me^{\bar{\omega}(t^*-t_0)}|y| + dK_5|z|(t^*-t_0)\bar{\rho}(N+\mu,y) +$$

$$+ \|Q_0 - P^{N+\mu}Q_0\| d(t^* - t_0) K_5 |z| \bar{M} e^{\bar{\omega}(t^* - t_0)} |y| =$$

$$= [f + d(t^* - t_0)][K_5 |z| \bar{\rho}(N+\mu, y) + K_5 \bar{M} e^{\bar{\omega}(t^* - t_0)} |y| |z| \|Q_0 - P^{N+\mu}Q_0\| +$$

$$+ \bar{M} e^{\bar{\omega}(t^* - t_0)} |y| [\rho(N+\mu, z) + \tilde{\rho}(N+\mu, Q^N) |z| K_6 + |\tilde{u} - \tilde{u}^{N,\mu}|_{L^2} k_5 + |z| \rho_Q(N) k_6]].$$

<u>Proof of Corollary 2.2</u>.

$$|J(t_0, P^{N+\mu}z, \tilde{u}^{N,\mu}) - J(t_0, z, \tilde{u})| =$$

$$= |\langle F^{N+\mu} z^{N,\mu}(t^*), z^{N,\mu}(t^*)\rangle - \langle \mathscr{F} z(t^*), z(t^*)\rangle| +$$

$$+ \int_{t_0}^{t^*} |\langle \mathscr{D}^{N+\mu}(t) z^{N,\mu}(t), z^{N,\mu}(t)\rangle - \langle \mathscr{D}(t) z(t), z(t)\rangle| dt +$$

$$+ \int_{t_0}^{t^*} |(C(t)\tilde{u}^{N,\mu}(t), \tilde{u}^{N,\mu}(t)) - (C(t)u(t), u(t))| dt \le$$

$$\le \langle P^{N+\mu} \mathscr{F} z^{N,\mu}(t^*) - \mathscr{F} z^{N,\mu}(t^*), z^{N,\mu}(t^*)\rangle +$$

$$+ (FP_1(z^{N,\mu}(t^*) - z(t^*)), P_1 z^{N,\mu}(t^*)) + (FP_1 z(t^*), P_1(z^{N,\mu}(t^*) - z(t^*))) +$$

$$+ \int_{t_0}^{t^*} |\langle \mathscr{D}^{N+\mu}(t) z^{N,\mu}(t) - \mathscr{D}(t) z^{N,\mu}(t), z^{N,\mu}(t)\rangle +$$

$$+ \langle \mathscr{D}(t)(z^{N,\mu}(t) - z(t)), z^{N,\mu}(t)\rangle + \langle \mathscr{D}(t) z(t), z^{N,\mu}(t) - z(t)\rangle| dt +$$

$$+ \int_{t_0}^{t^*} |(C(t)(\tilde{u}(t) - \tilde{u}^{N,\mu}(t)), \tilde{u}(t)) + (C(t)\tilde{u}^{N,\mu}(t), \tilde{u}(t) - \tilde{u}^{N,\mu}(t))| dt \le$$

$$= ((t^*-t_0)d+f)K_5|z|\{2[\rho(N+\mu,z) + \tilde{\rho}(N+\mu,Q^N)|z|K_6 + |\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}k_5 +$$

$$+ |z|\rho_Q(N)k_6] + ||P^{N+\mu}Q_0-Q_0||K_5|z|\} +$$

$$+ 2|\tilde{u}-\tilde{u}^{N,\mu}|_{L^2}c^{-1}K_1(t^*-t_0)^{1/2}\sup||C(t)||\ |z|,$$

which completes the proof.

Proof of Theorem 2.4. By Remarks 2.4 and (2.10) it follows that $\tilde{u}$ is continuous. Moreover by (2.15) we have

$$|\tilde{u}(t) - \tilde{u}^{N,0}(t)| = |C(t)^{-1}\mathcal{B}(t)^*(\Pi(t)S(t,t_0)z - \Pi^{N,0}(t)S^{N,0}(t,t_0)P^Nz)|.$$

Let $\varepsilon^N(t) = \operatorname{sgn} C(t)^{-1}\mathcal{B}(t)^*(\Pi(t)\ S(t,t_0)z - \Pi^{N,0}(t)S^{N,0}(t,t_0)P^Nz)$, then

$$|\tilde{u}(t) - \tilde{u}^{N,0}(t)| \leq \langle \Pi(t)S(t,t_0)z - \Pi^{N,0}S^{N,0}(t,t_0)P^Nz,\ \mathcal{B}(t)(C(t)^{-1})^*\varepsilon^N(t)\rangle.$$

By Corollary 2.1(c) and after our inspection of (2.16) and the proof of Theorem 2.3(b) it now follows that

$$|\tilde{u}(t) - \tilde{u}^{N,0}(t)| \leq \sup_{\xi} K_9[\bar{\rho}(N,\xi)|z| + \rho(N,z)|\xi| + |\xi||z|(\rho_Q(N) + \tilde{\rho}(N,Q_0))]$$

where the supremum is taken over $\{\mathcal{B}(t)(C(t)^{-1})^*\varepsilon^N(t) \mid t \in [t_0,t^*],\ N = 1,2,\ldots\}$ a relatively compact subset of Z. (H1), (H2), together with (vii) and (viii) in Section 2, and a simple compactness argument imply the result.

REFERENCES

[1] Banks, H.T., and J.A. Burns: Hereditary control problems: numerical methods based on averaging approximations, SIAM J. Control and Optimization 16, 169-208 (1978).

[2] Banks, H.T., J.A. Burns, E.M. Cliff and P.R. Thrift: Numerical solutions of hereditary control problems via an approximation technique, LCDS Technical Report 75-6.

[3] Banks, H.T. and F. Kappel: Spline approximations for functional differential equations, J. Differential Eqs. 34, 496-522 (1979).

[4] Delfour, M.C.: State theory of linear hereditary differential systems, J. Math. Anal. and Appl. 60, 8-35 (1977).

[5] Delfour, M.C.: The linear quadratic optimal control problem for hereditary differential systems: theory and numerical solutions, Appl. Mathematics and Optimization 3, 101-162 (1977).

[6] Gibson, J.S.: The Riccati integral equations for optimal control problems on Hilbert spaces, SIAM J. Control and Optimization 17, 537-565 (1979).

[7] Kappel, F., and K. Kunisch: Spline approximations for neutral functional differential equations, submitted to SIAM J. Numerical Analysis.

[8] Kappel, F., and W. Schappacher: Nonlinear functional differential equations and abstract integral equations, Proc. Royal Soc. Edinburgh 84A, 71-91 (1979).

[9] Kunisch, K. and W. Schappacher: Variation of constants formulas for partial differential equations with delay, to appear in: Nonlinear Analysis, Theory, Methods and Applications.

[10] Phillips, R.S.: The adjoint semigroup, Pacific J. Math. 5, 269-283 (1955).

[11] Prenter, P.M.: Splines and Variational Methods, Wiley Interscience (1975).

[12] Reber, D.C.: A finite difference technique for solving optimization problems governed by linear functional differential equations, J. Differential Eqs., 32, 192-232 (1979).

[13] Ross, D.W., and I. Flugge-Lotz: An optimal control problem for systems with differential-difference equation dynamics, SIAM J. Control, 7, 609-623 (1969).

[14] Schultz, M.H.: Spline Analysis, Prentice Hall (1973).

[15] Solimon, M.A. and W.H. Ray: Optimal feedback control for linear-quadratic systems having time delays, Int. J. Control, 15, 609-627 (1972).

END

DATE

FILMED

9-80

DTIC